

明 細 書

信号分離装置、信号分離方法、信号分離プログラム及び記録媒体
技術分野

[0001] 本発明は、信号処理の技術分野に属し、特に複数の源信号が空間内で混合された混合信号から源信号を抽出する技術に関する。

背景技術

[0002] 複数のセンサを用いることによりある信号を抽出してその他の信号を抑圧する従来の技術として、ビームフォーマ (beamformer) (ビームフォーミング (beamforming) と呼ぶ) が広く知られている (例えば、非特許文献1参照)。しかし、ビームフォーマでは、目的信号の方向に関する情報を必要とするため、そのような情報が得られない (または推定できない) 状況では利用しにくいという欠点がある。

また、より新しい別の技術としては、ブラインド信号分離 (BSS: Blind Signal Separation) が挙げられる (例えば、非特許文献2参照)。これは、上記ビームフォーマが必要とした情報を必要としないという点で優れており、様々な状況での利用が期待されている。以下、このブラインド信号分離による信号分離について説明する。

[0003] [ブラインド信号分離]

まず、ブラインド信号分離の定式化を行う。すべての信号はあるサンプリング周波数 f_s でサンプリングされ、離散的に表現されるものとする。また、 N 個の信号が混合されて M 個のセンサで観測されたとする。以下では、信号の発生源からセンサまでの距離により信号が減衰・遅延し、また壁などにより信号が反射して伝送路の歪みが発生する状況を扱う。このような状況で混合される信号は、信号源 k からセンサ q (q はセンサの番号を示す [$q=1, \dots, M$])。また、 k は信号源の番号を示す [$k=1, \dots, N$])。へのインパルス応答 $h_{qk}(r)$ による畳み込み混合

[0004] [数1]

$$x_q(t) = \sum_{k=1}^N \sum_{r=0}^{\infty} h_{qk}(r) s_k(t-r) \quad \dots (1)$$

で表現できる。ここで t はサンプリング時刻を示している。また、 $s_k(t)$ は、サンプリング

時刻 t において信号源から発せられる源信号を示しており、 $x_q(t)$ は、サンプリング時刻 t においてセンサ q で観測される信号を示している。また、 r は掃引のための変数を示している。

一般的なインパルス応答 $h_{qk}(r)$ は、適当な時間経過後にパルスの強い応答を持ち、時間と共に減衰していく。ブラインド信号分離の目的は、源信号 $s_1(t), \dots, s_N(t)$ やインパルス応答 $h_{11}(r), \dots, h_{1N}(r), \dots, h_{M1}(r), \dots, h_{MN}(r)$ を知らずに、観測信号(以下「混合信号」と呼ぶ) $x_1(t), \dots, x_M(t)$ のみから、源信号 $s_1(t), \dots, s_N(t)$ にそれぞれ対応する分離信号 $y_1(t), \dots, y_N(t)$ を求めることにある。

[0005] [周波数領域]

次に、従来のブラインド信号分離の手順について説明する。

ここでは周波数領域において分離の操作を行う。そのためにセンサ q での混合信号 $x_q(t)$ に L 点の短時間離散フーリエ変換(STFT: Short-Time Fourier Transform)を適用し、周波数ごとの時間系列

[0006] [数2]

$$X_q(f, \tau) = \sum_{r=-L/2}^{(L/2)-1} x_q(\tau+r) g(r) e^{-j2\pi f r} \quad \dots (2)$$

を求める。ここで f は周波数であり $f=0, f_s/L, \dots, f_s(L-1)/L$ と離散化されている(f_s はサンプリング周波数)。また、 τ は離散時間であり、 j は虚数単位である。さらに $g(r)$ は窓関数である。窓関数としては、例えば、ハニング窓

[0007] [数3]

$$g(r) = \frac{1}{2} \left(1 + \cos \frac{2\pi r}{L} \right)$$

などの $g(0)$ にパワーの中心を持つ窓関数を用いる。この場合、 $X_q(f, \tau)$ は時刻 $t=\tau$ を中心とする混合信号 $x_q(t)$ の周波数特性を表現する。なお、 $X_q(f, \tau)$ は L サンプルにわたる情報を含んでいるため、すべての τ に対して $X_q(f, \tau)$ を求める必要はなく、適当な間隔の τ ごとに $X_q(f, \tau)$ を求める。

周波数領域で処理を行うと、式(1)で示される時間領域での畳み込み混合が、

[0008] [数4]

$$X_q(f, \tau) = \sum_{k=1}^N H_{qk}(f) S_k(f, \tau) \quad \cdots (3)$$

と各周波数での単純混合に近似表現でき、分離の操作が単純になる。ここで、 $H_{qk}(f)$ は源信号 k からセンサ q までの周波数応答であり、 $S_k(f, \tau)$ は式(2)と同様な式に従って源信号 $s_k(t)$ に短時間離散フーリエ変換を施したものである。ベクトルを用いて式(3)を表記すると、

[0009] [数5]

$$X(f, \tau) = \sum_{k=1}^N H_k(f) S_k(f, \tau) \quad \cdots (4)$$

となる。ここで、 $X(f, \tau) = [X_1(f, \tau), \dots, X_M(f, \tau)]^T$ は混合信号ベクトル、 $H_k(f) = [H_{1k}(f), \dots, H_{Mk}(f)]^T$ は信号源 k から各センサへの周波数応答をまとめたベクトルである。なお、 $[*]^T$ は $[*]$ の転置ベクトルを示す。

[独立成分分析による信号分離]

ブラインド信号分離手法の1つとして、独立成分分析(ICA:Independent Component Analysis)による信号分離がある。このICAによる手法では、混合信号ベクトル $X(f, \tau)$ のみから、 N 行 M 列の分離行列 $W(f)$ 及び分離信号ベクトル

$$Y(f, \tau) = W(f) X(f, \tau) \quad \cdots (5)$$

を算出する。ここで、分離行列 $W(f)$ は、分離信号ベクトル $Y(f, \tau) = [Y_1(f, \tau), \dots, Y_N(f, \tau)]^T$ の各要素(分離信号) $Y_1(f, \tau), \dots, Y_N(f, \tau)$ が互いに独立になるように算出される。そのためのアルゴリズムには、非特許文献4に記載されているものなどがある。

[0010] ICAでは信号の独立性に着目して分離を行うため、得られる分離信号 $Y_1(f, \tau), \dots, Y_N(f, \tau)$ には、順序の任意性がある。信号の順序が入れ替わってもそれらの独立性は保たれるからである。この順序の任意性の問題は、パーミュテーション(permutation)の問題と呼ばれ、周波数領域での信号分離において非常に重要な問題である。そして、このパーミュテーションの問題は、同じ源信号 $S_k(f, \tau)$ に対応する分離信号 $Y_p(f, \tau)$ の添字 p が、すべての周波数 f で同じになるように解決されなければならない。

このパーミュテーション問題を解決する従来手法として、例えば非特許文献5に示されるものがある。この手法では、選択された2つのセンサ(センサペア)の位置を基準として信号源の位置に関する情報(方向と距離比)を推定する。そして、複数のセンサペアでの推定値を統合することで、より詳細な位置情報が得られる。そして、これらの位置情報としての推定値をクラスタリングし、同じクラスに属する推定値は同じ信号源に対応するものとみなしてパーミュテーション問題を解決する。

[0011] [時間周波数マスクによる信号分離]

また、他のブラインド信号分離手法として、時間周波数マスクによる方法がある。この手法は、信号源の数 N とセンサ数 M が $M < N$ の関係にある場合でも有効な信号分離抽出手法である。

この手法では信号のスパース性を仮定する。スパースとは、信号が殆どの離散時間 τ において0であることを指す。信号のスパース性は、例えば、周波数領域での音声信号で確認される。信号のスパース性と相互独立性を仮定することで、複数の信号が同時に存在していても、各時間周波数ポイント (f, τ) では互いに重なって観測される確率が低いことを仮定できる。よって、各時間周波数ポイント (f, τ) の各センサにおける混合信号は、その時間周波数ポイント (f, τ) でアクティブな1つの信号 $s_p(f, \tau)$ のみから成ると仮定できる。よって、混合信号ベクトルを適当な特徴量によりクラスタリングし、各クラス C_k のメンバの時間周波数 (f, τ) に対応する混合信号 $X(f, \tau)$ を抽出する時間周波数マスク $M_k(f, \tau)$ を推定し、

$$Y_k(f, \tau) = M_k(f, \tau) X_{Q'}(f, \tau)$$

により、各信号を分離抽出する。ここで、 $X_{Q'}(f, \tau)$ は混合信号のうちの1つであり、 $Q' \in \{1, \dots, M\}$ である。

- [0012] クラスタリングに用いる特徴量としては、例えば、2つのセンサ(センサ q と基準センサ Q [なお、 Q を基準値と呼び、基準値 Q に対応するセンサを基準センサ Q と表記する。])における混合信号の位相差

[0013] [数6]

$$\phi(f, m) = \angle \frac{X_q(f, \tau)}{X_Q(f, \tau)} \quad \dots (8)$$

から計算される信号の推定到来方向 (Direction of Arrival : DOA)

[0014] [数7]

$$\theta(f, \tau) = \cos^{-1} \frac{\phi(f, \tau) \cdot c}{2\pi \cdot f \cdot d} \quad \dots (9)$$

を例示できる(例えば、非特許文献3参照)。なお、 d はセンサ q と基準センサ Q との距離であり、 c は信号速度である。また、クラスタリングには、 k -means法(例えば、非特許文献6参照)等を用いることができる。また、時間周波数マスク $M_k(f, \tau)$ としては、例えばそれぞれのクラスタ C_k に属するメンバの平均値 $\theta_1 \sim \theta_2 \sim \dots, \theta_N \sim$ を求め、

[0015] [数8]

$$M_k(f, \tau) = \begin{cases} 1 & \theta_k \sim - \Delta \leq \theta(f, \tau) \leq \theta_k \sim + \Delta \\ 0 & \text{otherwise} \end{cases} \quad (k=1, \dots, N)$$

のようにして生成したものを用いることができる。ここで Δ は信号を抽出する範囲を与える。この方法では、 Δ を小さくすると、よい分離抽出性能が得られるが非線形型歪みは大きくなる。また、 Δ を大きくすると、非線形型歪みは減少するが分離性能が劣化する。

その他、クラスタリングの特徴量として、2つのセンサ(センサ q と基準センサ Q)における混合信号の位相差(式(8))や両者のゲイン比

[0016] [数9]

$$\alpha(f, \tau) = \left| \frac{X_q(f, \tau)}{X_Q(f, \tau)} \right|$$

を用いてもよい。

非特許文献1: B. D. Van Veen and K. M. Buckley, "Beamforming: a versatile approach to spacial filtering," IEEE ASSP Magazine, pp. 4-24, April 1988

非特許文献2: S. Haykin, eds, "Unsupervised Adaptive Filtering," John-Wiley & Sons, 2000, ISBN 0-471-29412-8

非特許文献3: S. Araki, S. Makino, A. Blin, R. Mukai, and H. Sawada, "Underdetermined blind separation for speech in real environments with sparseness and ICA," in P

roc. ICASSP 2004, vol. III, May 2004, pp. 881-884

非特許文献4: A. Hyvarinen and J. Karhunen and E. Oja, "Independent Component Analysis," John Wiley & Sons, 2001, ISBN 0-471-40540

非特許文献5: R. Mukai, H. Sawada, S. Araki and S. Makino, "Frequency Domain Blind Source Separation using Small and Large Spacing Sensor Pairs," in Proc. of ICASSP 2004, vol. V, pp.1-4, May 2004.

非特許文献6: R. O. Duda, P. E. Hart, and D. G. Stork, Pattern Classification, Wiley Interscience, 2nd edition, 2000

発明の開示

発明が解決しようとする課題

[0017] しかし、上述した従来の技術では、複数のセンサによる観測信号から得られた情報を容易かつ効率的に利用して信号分離処理を行うことができないという問題点があった。

例えば、独立成分分析による信号分離の場合、正確にパーミュテーション問題を解決するためには複雑な操作が必要となるという問題がある。すなわち、従来のパーミュテーション問題の解決手法では、センサペア毎に方向と距離比とが推定される。従って、正確にパーミュテーション問題を解決するためには複数のセンサペアでの推定値を統合する必要があった。また、この推定値には誤差が含まれているが、より誤差が少ないと思われるセンサペアを優先的に用いたり、誤差をうまく吸収できるように統合方法を工夫したりする必要があった。さらに、この従来手法では、信号源の位置に関する情報を推定する必要性から、センサの位置情報をあらかじめ取得しておかなければならないという問題点もあった。これはセンサを不規則に配置する場合に不利である。また、たとえセンサを規則的に配置するとしても、配置情報を厳密に知ることは困難であり、より正確なパーミュテーション問題の解決のためにはキャリブレーションなどの操作を必要とする。

[0018] また、従来の時間周波数マスクによる信号分離手法の場合、センサが2個の場合のための手法しか提案されておらず、センサが3個以上ある場合であっても、そのうち特定の2つのセンサ q , Q のみの情報しか用いずに特徴量を計算していた。これは次

元の低下を意味し、全てのセンサを用いる場合に比べて情報量が欠落する。そのため、全てのセンサの情報を効率的に使えず性能に限度があった。また、全てのセンサの情報を有効利用するには、例えば、非特許文献5にならって、複数のセンサペアで求めた特徴量を統合することも可能であろうが、この統合のためには、特徴量抽出のための更なる処理が必要であり、より統合誤差が少ないと思われるセンサペアを優先的に用いるなど、統合の際に工夫する必要があると思われる。さらに、この手法でも、センサの厳密な配置情報をあらかじめ知っておかなければならないという問題点がある。これは、センサを自由に配置する場合に不利である。また、たとえ規則的に配置するとしても、配置情報を厳密に知ることは困難であり、より正確な信号抽出のためにはキャリブレーションなどの操作を必要とする。

[0019] さらに、ブラインド信号分離の処理の基本は、センサで観測された混合信号を分離して、複数の分離信号を出力することである。しかし、すべての分離信号が重要なわけではなく、一部の分離信号のみに目的とする目的信号が含まれていることもある。そのような場合、目的信号が含まれている分離信号を選択する必要があるが、従来のブラインド信号分離では、どの分離信号に目的信号が含まれているかという情報までは提供されない。従って、何らかの別の手段によって、どの分離信号に目的信号が含まれているかを判定しなければならない。

[0020] 本発明はこのような点に鑑みてなされたものであり、複数のセンサによる観測信号から得られた情報を容易かつ効率的に利用して信号分離処理を行うことが可能な技術を提供することを目的とする。

課題を解決するための手段

[0021] 本発明では、上記課題を解決するために、まず、周波数領域変換部が、複数のセンサで観測された混合信号を、周波数領域の混合信号に変換する。次に、正規化部が、当該周波数領域の混合信号を用いて生成された複素ベクトルを正規化し、当該複素ベクトルの周波数依存性を排除した正規化ベクトルを生成する。次に、クラスタリング部が、当該正規化ベクトルをクラスタリングし、クラスタを生成する。そして、このクラスタを用いて信号分離処理を行う。

ここで、当該クラスタの生成には、混合信号を観測する各センサの厳密な位置情報

を、入力情報として直接用いることは必要としない。また、各クラスは、各信号源の位置に依存する情報に基づき形成される。これらより、本発明では、各センサの厳密な位置情報を用いることなく、信号分離処理が可能となる。

- [0022] また、本発明において好ましくは、正規化部は、複素ベクトルが具備する特定の1つの要素を基準として当該複素ベクトルの各要素の偏角を正規化する第1正規化部と、第1正規化部で正規化された各要素の偏角を周波数に比例した値で除算する第2正規化部とを有する。

これらの正規化後の複素ベクトルは、信号源の位置に依存したクラスを形成する。これにより、各センサの厳密な位置情報を用いることなく、信号分離処理を行うことが可能となる。

- [0023] また、本発明において好ましくは、正規化部は、第2正規化部で正規化された各要素からなるベクトルのノルムを所定の値に正規化する第3正規化部をさらに有する。

これらの正規化後の複素ベクトルは、信号源の位置に依存したクラスを形成する。また、第2正規化された各要素からなるベクトルのノルムを正規化することにより、クラスタリングの処理が簡略化される。

また、好ましい第1の本発明では、まず、周波数領域変換部が、複数のセンサで観測された混合信号を、周波数領域の混合信号に変換する。次に、分離行列算出部が、周波数領域の混合信号を用いて、周波数毎に分離行列を算出し、逆行列算出部が、分離行列の一般化逆行列を算出する。次に、基底ベクトル正規化部が、一般化逆行列を構成する基底ベクトルの正規化を行い、正規化基底ベクトルを算出する。次に、クラスタリング部が、正規化基底ベクトルをクラスタリングし、クラスを生成する。その後、順列算出部が、クラスタの中心ベクトルと正規化基底ベクトルとを用い、分離行列の要素を並び替えるための順列を算出する。なお、基底ベクトルは、複素ベクトルに含まれる概念である。

- [0024] この第1の本発明では、基底ベクトルを正規化してクラスタリングし、パーミュテーション問題解決のための順列を算出する。そのため、このクラスタリングに際し、センサの位置情報をあらかじめ取得しておく必要がない。さらに、この好ましい本発明では、正規化基底ベクトルのすべての要素を対象としてクラスタリングを行い、その結果から

パーミュテーション問題解決のための順列を算出する。そのため、従来のように推定結果を統合するための操作を必要としない。

また、当該第1の本発明において、より好ましくは、基底ベクトル正規化部は、基底ベクトルの周波数依存性を排除する正規化を行う。そして、より好ましくは、当該基底ベクトルの周波数依存性を排除する正規化は、基底ベクトルが具備する特定の1つの要素を基準として当該基底ベクトルの各要素の偏角を正規化し、さらに、各要素の偏角を周波数に比例した値で除算する正規化である。このような正規化により、信号源の位置に依存したクラスタの形成が可能となる。

[0025] また、当該第1の本発明において、より好ましくは、基底ベクトルの周波数依存性を排除する正規化は、基底ベクトル $A_p(f)$ ($p=1, \dots, N$, N は信号源の数)の各要素 $A_{qp}(f)$ ($q=1, \dots, M$, M は混合信号を観測するセンサの数)に対し、

[0026] [数10]

$$A_{qp}'(f) = |A_{qp}(f)| \exp \left[j \frac{\arg[A_{qp}(f)/A_{Qp}(f)]}{4fc^{-1}d} \right] \quad \dots (10)$$

の演算を行う正規化である。ただし、 \exp をネピア数とし、 $\arg[\cdot]$ を偏角とし、 f を周波数とし、 j を虚数単位とし、 c を信号の伝達速度とし、 Q を M 以下の自然数から選択された基準値とし、 d を実数とする。すなわち、式(10)の演算によって行われる正規化は、基底ベクトルが具備する特定の1つの要素を基準として当該基底ベクトルの各要素の偏角を正規化し、さらに、各要素の偏角を周波数に比例した値で除算する正規化である。この正規化により、周波数への依存性を排除できる。また、当該正規化には、センサの厳密な配置情報を必要としない。

[0027] また、上述の式(10)の実数 d は、要素 $A_{Qp}(f)$ に対応する基準センサ Q と他のセンサとの最大距離 d_{\max} であることが望ましい。クラスタリングの精度が一般的に向上するからである。この詳細は後述する。

また、第1の本発明において好ましくは、基底ベクトルを周波数に依存しない周波数正規化ベクトルに正規化し、この周波数正規化ベクトルをノルムが規定値をとる正規化基底ベクトルに正規化する。以上の二段階の正規化によって生成された正規化基底ベクトルは、周波数に依存せず、信号源の位置のみに依存する。なお、ノルムの

正規化により、クラスタリングの処理が簡略化できる。

- [0028] さらにまた、第1の本発明において好ましくは、分離信号のエンベロープ(分離信号の絶対値の包絡線)とクラスタの中心ベクトルと正規化基底ベクトルとを用いて順列を算出する。これにより、より精度よくパーミュテーション問題を解決することができる。

また、好ましい第2の本発明では、周波数領域変換部が、複数のセンサで観測された混合信号を周波数領域の混合信号に変換し、信号分離部が、周波数領域の混合信号を用い、周波数毎に分離行列と分離信号とを算出する。そして、目的信号選択部が、複数の分離信号から、目的信号を含むものを選択信号として選択する。その手順は、分離行列の一般化逆行列の列である基底ベクトルを正規化し、その正規化された基底ベクトルをクラスタリングし、そのクラスタの分散を指標として選択信号を決定するものである。なお、分離行列が正方行列の場合、「一般化逆行列」は逆行列に相当する。すなわち、ここでの一般化逆行列は、通常の逆行列をも含む概念である。

- [0029] クラスタの分散を指標とすることで、センサに近い信号を目的信号として位置付け、そのような目的信号を含む分離信号を選択信号として選り出すことができる。その理由を以下に述べる。ここで、基底ベクトルの正規化は、複数の信号源から発せられた信号の畳み込み混合を所定のモデル(例えば、近距離場モデル)に近似した場合に、正規化された基底ベクトルが信号源の位置のみに依存したクラスタを形成するように行われる。しかし、実環境では、このようなモデルに反映されない様々な因子が存在する。例えば、近距離場モデルでは、信号が壁などに反射することによって生じる伝送歪みが反映されない。このような実環境とモデルとの乖離は、信号源からセンサまでの距離が大きいほど大きくなり、センサに近い信号ほど小さくなる。従って、センサに近い信号ほど実環境に近い条件で正規化ができ、実環境とモデルとの乖離に起因するクラスタの分散を小さくできる。好ましい第2の本発明では、この関係に着目し、クラスタの分散の大きさを指標として、センサに近い信号である目的信号を含む選択信号を抽出する。以上の操作により、目的信号を抽出し、他の妨害信号をある程度抑圧することができる。

- [0030] しかし、分離行列及び分離信号の算出を独立成分分析(ICA:Independent Component Analysis)で行う場合、以上の処理によって完全に抑制できる妨害信号の数は

センサの数から1を引いた数までである。すなわち、妨害信号の数がそれ以上の場合には抑圧しきれない残留成分が残る。そのため、本発明において好ましくは、更に、マスク生成部が、周波数領域の混合信号と基底ベクトルとを用いて時間周波数マスクを生成し、マスキング部が、この時間周波数マスクを、選択された選択信号に適用する。これにより、信号源の数がセンサ数より多い場合でも、選択信号に残留した妨害信号をより良く抑圧することができる。

[0031] 更に、第2の本発明において好ましくは、マスク生成部は、周波数領域の混合信号を用いて白色化行列を生成し、白色化行列を用い、周波数領域の混合信号を要素とする混合信号ベクトルを変換した白色化混合信号ベクトルと、基底ベクトルを変換した白色化基底ベクトルとを算出し、白色化混合信号ベクトルと白色化基底ベクトルとがなす角度を時間周波数毎に算出し、この角度を要素とする関数を用いて時間周波数マスクを生成する。この時間周波数マスクを選択信号に適用することにより、選択信号に残留した妨害信号を更に抑圧できる。

[0032] また更に、第2の本発明において好ましくは、白色化行列が、周波数を f とし、離散時間を τ とし、混合信号ベクトルを $X(f, \tau)$ とし、ベクトル $*$ を時間平均したベクトルを $\langle * \rangle$ とし、ベクトル $*$ の複素共役転置ベクトル(各要素の複素共役を採り転置したベクトル)を $*^H$ とし、 $R(f) = \langle X(f, \tau) \cdot X(f, \tau)^H \rangle$ とした場合における $V(f) = R(f)^{-1/2}$ である。そして、 $Z(f, \tau) = V(f) \cdot X(f, \tau)$ の演算により、白色化混合信号ベクトル $Z(f, \tau)$ を算出し、基底ベクトルを $A(f)$ とした場合における $B(f) = V(f) \cdot A(f)$ の演算により、白色化基底ベクトル $B(f)$ を算出する。また、ベクトル $*$ の絶対値を $|*|$ とし、ベクトル $*$ のノルムを $\|*\|$ とした場合における $\theta(f, \tau) = \cos^{-1}(|B^H(f) \cdot Z(f, \tau)| / \|B(f)\| \cdot \|Z(f, \tau)\|)$ の演算により、角度 $\theta(f, \tau)$ を算出し、 α, g, θ_T を実数とした場合におけるロジスティック関数 $M(\theta(f, \tau)) = \alpha / (1 + e^{g(\theta(f, \tau) - \theta_T)})$ を時間周波数マスクとして算出する。この時間周波数マスクを、抽出した選択信号に適用することにより、選択信号に残留した妨害信号を更に抑圧できる。

[0033] また、第2の本発明において好ましくは、目的信号選択部は、基底ベクトルの周波数依存性を排除する正規化を行う。第2の本発明において、より好ましくは、基底ベクトルの周波数依存性を排除する正規化は、基底ベクトルが具備する特定の1つの要

素を基準 s として当該基底ベクトルの各要素の偏角を正規化し、さらに、各要素の偏角を周波数に比例した値で除算する正規化である。またさらに、第2の本発明において、好ましくは、基底ベクトルの周波数依存性を排除する正規化は、基底ベクトル $A_p(f)$ (p は自然数)の各要素 $A_{qp}(f)$ ($q=1, \dots, M$ 、 M は混合信号を観測するセンサの数)に対し、

[0034] [数11]

$$A_{qp}'(f) = |A_{qp}(f)| \exp \left[j \frac{\arg[A_{qp}(f)/A_{Qp}(f)]}{4fc^{-1}d} \right] \quad \dots (11)$$

の演算を行う正規化である。ただし、 \exp をネピア数とし、 $\arg[\cdot]$ を偏角とし、 f を周波数とし、 j を虚数単位とし、 c を信号の伝達速度とし、 Q を M 以下の自然数から選択された基準値とし、 d を実数とする。このような正規化により、複数の信号源から発せられた信号の畳み込み混合を所定のモデルに近似した場合に、正規化された基底ベクトルが信号源の位置に依存したクラスタを形成することとなる。その結果、上述のようにクラスタの分散の大きさを指標として、目的信号が含まれる分離信号を選択することが可能となる。なお、当該正規化には、センサの厳密な配置情報を必要としない。

[0035] またさらに、上述の式(11)の実数 d は、基準センサ Q と他のセンサとの最大距離 d_{\max} であることが望ましい。クラスタリングの精度が一般的に向上するからである。この詳細は後述する。

また、第2の本発明において好ましくは、目的信号選択部は、分散が最小となるクラスタを選択し、選択されたクラスタに対応する分離信号を目的信号が含まれるものとして選択する。これにより、モデルからの乖離が最も少ない信号(例えば、センサに最も近い信号)を目的信号として抽出することができる。

[0036] また、好ましい第3の本発明では、まず、周波数領域変換部が、複数のセンサで観測された混合信号を周波数領域の混合信号に変換する。次に、ベクトル正規化部が、周波数領域の混合信号から構成される混合信号ベクトルの正規化を行い、正規化ベクトルを算出する。そして、クラスタリング部が、正規化ベクトルをクラスタリングし、クラスタを生成する。その後、分離信号生成部が、 k 番目のクラスタに属する正規化ベク

トルの時間周波数に対応する混合信号ベクトルから所定番目の要素を抽出し、これをk番目の要素とした分離信号ベクトルを生成する。

[0037] ここで、第3の本発明では、全てのセンサで観測された混合信号を正規化してクラスタリングを行い、各クラスタの情報を用いて分離信号ベクトルを生成している。これは、全てのセンサの情報を同時に用いて分離信号を抽出していることを意味する。また、この処理では、センサの厳密な配置情報を必要としない。以上より、第3の本発明では、センサの厳密な配置情報を必要とすることなく、全ての観測信号から得られた情報を容易かつ効率的に利用して信号分離を行うことができる。

また、第3の本発明において、好ましくは、ベクトル正規化部は、周波数領域の混合信号から構成される混合信号ベクトルの周波数依存性を排除する正規化を行う。また、さらに好ましくは、混合信号ベクトルの周波数依存性を排除する正規化は、混合信号ベクトルが具備する特定の1つの要素を基準として当該基底ベクトルの各要素の偏角を正規化し、さらに、各要素の偏角を周波数に比例した値で除算する正規化である。またさらに好ましくは、混合信号ベクトルの周波数依存性を排除する正規化は、混合信号ベクトルの各要素 $X_q(f, \tau)$ ($q=1, \dots, M$, M は混合信号を観測するセンサの数)に対し、

[0038] [数12]

$$X_q'(f, \tau) = |X_q(f, \tau)| \exp \left[j \frac{\arg[X_q(f, \tau)/X_Q(f, \tau)]}{4fc^{-1}d} \right] \quad \dots (12)$$

の演算を行う正規化である。ただし、 \exp をネピア数とし、 $\arg[\cdot]$ を偏角とし、 j を虚数単位とし、 c を信号の伝達速度とし、 Q を M 以下の自然数から選択された値とし、 d を実数とし、 f を周波数とし、 τ を離散時間とする。これにより、周波数の依存性を排除できる。これにより、信号源の位置に依存したクラスタを形成できる。なお、当該正規化には、センサの厳密な配置情報を必要としない。

[0039] また、上述の式(12)の実数 d は、要素 $X_q(f, \tau)$ に対応するセンサと他のセンサとの最大距離 d_{\max} であることが望ましい。クラスタリングの精度が一般的に向上するからである。この詳細は後述する。

また、第3の本発明において、好ましくは、ベクトル正規化部は、混合信号ベクトル

の周波数依存性を排除する正規化と、ノルムを規定値にする正規化とを行う。これにより、クラスタリングの処理が簡略化される。

発明の効果

[0040] 以上のように、本発明では、複数のセンサによる観測信号から得られた情報を容易かつ効率的に利用して信号分離処理を行うことができる。

例えば、第1の発明では、厳密なセンサ配置に係る情報の事前取得や煩雑な操作を必要とせず正確にパーミュテーション問題を解決できる。また、第2の発明では、目的信号の方向に関する情報を知ることなく、複数の信号源から発せられた信号が混合された混合信号から目的信号を抽出することができる(たとえ $N > M$ であったとしても)。また、第3の発明では、センサの厳密な配置情報を必要とすることなく、全ての観測信号から得られた情報を容易かつ効率的に利用して信号分離を行うことができる(たとえ $N > M$ であったとしても)。

図面の簡単な説明

[0041] [図1]図1は、本発明の原理を備えた信号分離装置の機能構成を例示したブロック図である。

[図2]図2は、第1の実施の形態における信号分離装置のハードウェア構成を例示したブロック図である。

[図3]図3は、第1の実施の形態における信号分離装置のブロック図の例示である。

[図4]図4Aは、図3におけるパーミュテーション問題解決部の詳細を例示したブロック図である。図4Bは、図4Aの基底ベクトル正規化部の詳細を例示したブロック図である。

[図5]図5は、第1の実施の形態における信号分離装置の処理の全体を説明するためのフローチャートである。

[図6]図6は、第1の実施の形態の正規化処理の詳細を説明するためのフローチャートである。

[図7]図7Aは、 $d_{\max} / 2 \geq d$ である場合における、パラメータ d 毎の正規化基底ベクトルの要素 $A_{qp}''(\hat{r})$ とその偏角 $\arg[A_{qp}''(\hat{r})]$ との関係を説明するための複素平面図である。図7Bは、 $d_{\max} / 2 < d < d_{\max}$ である場合における、パラメータ d 毎の正規化基底ベクトル

ルの要素 $A''(f)$ とその偏角 $\arg[A''(f)]$ との関係を説明するための複素平面図である。

[図8]図8Aは、 $d=d_{\max}$ である場合における、パラメータ d 毎の正規化基底ベクトルの要素 $A''(f)$ とその偏角 $\arg[A''(f)]$ との関係を説明するための複素平面図である。

図8Bは、 $d>d_{\max}$ である場合における、パラメータ d 毎の正規化基底ベクトルの要素 $A''(f)$ とその偏角 $\arg[A''(f)]$ との関係を説明するための複素平面図である。

[図9]図9は、第2の実施の形態における信号分離装置のブロック図の例示である。

[図10]図10Aは、図9におけるパーミュテーション問題解決部の詳細を例示したブロック図である。図10Bは、図10Aの順列修正部の詳細を例示したブロック図である。

[図11]図11は、第2の実施の形態における信号分離装置の処理の全体を説明するためのフローチャートである。

[図12]図12は、図11におけるステップS58の例を説明するためのフローチャートである。

[図13]図13は、図11におけるステップS58の例を説明するためのフローチャートである。

[図14]図14Aは、第1の実施の形態及び第2の実施の形態による音源分離実験条件を示した図である。図14Bは、第1の実施の形態及び第2の実施の形態による音源分離実験結果を示した図である。

[図15]図15Aは、第1の実施の形態及び第2の実施の形態による音源分離実験条件を示した図である。図15Bは、第1の実施の形態及び第2の実施の形態による音源分離実験結果を示した図である。

[図16]図16は、第3の実施の形態における信号分離装置の構成を例示したブロック図である。

[図17]図17Aは、図16における目的信号選択部の詳細構成を例示したブロック図である。図17Bは、図17Aにおける基底ベクトルクラスタリング部の詳細構成を例示したブロック図である。

[図18]図18Aは、図16における時間周波数マスキング部の詳細構成を例示したブロック図である。図18Bは、図18Aにおけるマスク生成部の詳細を例示したブロック図

である。

[図19]図19は、第3の実施の形態における信号分離処理の全体を説明するためのフローチャートである。

[図20]図20は、第3の実施の形態における目的信号選択部の処理の詳細を説明するためのフローチャートである。

[図21]図21Aは、ステップS112の周波数正規化の詳細を説明するためのフローチャートである。図21Bは、ステップS113のノルム正規化の詳細を説明するためのフローチャートである。

[図22]図22は、選択信号の選択手順(ステップS115)の詳細を例示したフローチャートである。

[図23]図23は、図19におけるステップS104の詳細を説明するためのフローチャートである。

[図24]図24Aは、二種類の実数パラメータ θ_T, g に対し、式(46)に従って算出された時間周波数マスク $M(f, \tau)$ を例示した図である。図24Bは、ある時間周波数位置 (f, τ) において目的信号 $(I(f)=1)$ に対するベクトル $V(f) \cdot H_1(f)$ の他に、妨害信号に対応ベクトルする $V(f) \cdot H_2(f), V(f) \cdot H_3(f)$ が共存することを示した図である。

[図25]図25は、第4の実施の形態における信号分離装置のブロック図の例示である。

[図26]図26は、第4の実施の形態における信号分離装置の処理を説明するためのフローチャートである。

[図27]図27は、第5の実施の形態における信号分離装置のブロック図の例示である。

[図28]図28Aは、図27における時間周波数マスク生成部の詳細構成を示すブロック図である。図28Bは、図28Aのマスク生成部の詳細構成を示すブロック図である。

[図29]図29は、第5の実施の形態における時間周波数マスクの生成処理を説明するためのフローチャートである。

[図30]図30Aは、図29におけるステップS171の詳細を説明するためのフローチャートである。図30Bは、図29におけるステップS172の詳細を説明するためのフロー

チャートである。

[図31]図31Aは、第3の実施の形態及び第4の実施の形態による効果を示すための実験条件を示した図である。図31Bは、ICAのみの場合(第4の実施の形態)と、ICAと時間周波数マスキングの両方を組み合わせた場合(第3の実施の形態)のSIRの平均改善量を示す表である。

[図32]図32は、第6の実施の形態における信号分離装置のブロック図の例示である。

[図33]図33は、図32における信号分離部の詳細を例示したブロック図である。

[図34]図34は、第6の実施の形態における信号分離装置の処理の全体を説明するためのフローチャートである。

[図35]図35Aは、図34に示したステップS202の処理の詳細を説明するためのフローチャートである。図35Bは、図34に示したステップS203の処理の詳細を説明するためのフローチャートである。

[図36]図36は、図34に示したステップS205の処理の詳細を説明するためのフローチャートである。

[図37]図37Aは、 $d_{\max}/2 \geq d$ の場合における、パラメータ d 毎のノルム正規化ベクトル $X''(f, \tau)$ の要素 $X''_q(f, \tau)$ と、その偏角 $\arg[X''_q(f, \tau)]$ との関係を説明するための複素平面図である。図37Bは、 $d_{\max}/2 < d < d_{\max}$ の場合における、パラメータ d 毎のノルム正規化ベクトル $X''(f, \tau)$ の要素 $X''_q(f, \tau)$ と、その偏角 $\arg[X''_q(f, \tau)]$ との関係を説明するための複素平面図である。

[図38]図38Aは、 $d = d_{\max}$ の場合における、パラメータ d 毎のノルム正規化ベクトル $X''(f, \tau)$ の要素 $X''_q(f, \tau)$ と、その偏角 $\arg[X''_q(f, \tau)]$ との関係を説明するための複素平面図である。図38Bは、 $d > d_{\max}$ の場合における、パラメータ d 毎のノルム正規化ベクトル $X''(f, \tau)$ の要素 $X''_q(f, \tau)$ と、その偏角 $\arg[X''_q(f, \tau)]$ との関係を説明するための複素平面図である。

[図39]図39Aは、第6の実施の形態による音源分離実験条件を示した図である。図39Bは、第6の実施の形態による音源分離実験結果を示した図である。

[図40]図40Aは、第6の実施の形態による音源分離実験条件を示した図である。図4

0Bは、第6の実施の形態による音源分離実験結果を示した図である。

[図41]図41Aは、第6の実施の形態による音源分離実験条件を示した図である。図41Bは、第6の実施の形態による音源分離実験結果を示した図である。

符号の説明

[0042] 1, 10, 200, 1001, 1200, 1300, 2001 信号分離装置

発明を実施するための最良の形態

[0043] 以下、本発明の実施の形態を図面を参照して説明する。

〔原理〕

まず、本発明の原理について説明する。

図1は、本発明の原理を備えた信号分離装置1の機能構成を例示したブロック図である。なお、後述のように、信号分離装置1は、例えば、公知のノイマン型のコンピュータに所定のプログラムを実行させることにより構成される。

信号分離装置1は、複数の信号源から発せられた源信号の混合からなる混合信号を当該源信号に分離する装置である。図1に例示するように、信号分離装置1は、周波数領域変換部2と、複素ベクトル生成部3と、正規化部4と、クラスタリング部5とを有している。ここで、正規化部4は、複素ベクトルが具備する特定の1つの要素を基準として当該複素ベクトルの各要素の偏角を正規化する第1正規化部4aと、第1正規化部4aで正規化された各要素の偏角を周波数に比例した値で除算する第2正規化部4bと、第2正規化部4bで正規化された各要素からなるベクトルのノルムを所定の値に正規化する第3正規化部4cとを有する。なお、第1正規化部4a及び第2正規化部4bにより、複素ベクトルの周波数依存性が排除される(周波数正規化)。

[0044] 信号分離装置1によって信号分離処理を行う場合、まず、複数のセンサで観測された混合信号(時間領域の信号)が周波数領域変換部2に入力される。周波数領域変換部2は、短時間離散フーリエ変換等によって、複数のセンサで観測された混合信号(時間領域の信号)を、周波数領域の混合信号に変換する。次に、複素ベクトル生成部3が、当該周波数領域の混合信号を用い、複素数の要素からなる複素ベクトルを生成する。次に、正規化部4が、当該複素ベクトルを正規化し、当該複素ベクトルの周波数依存性を排除した正規化ベクトルを生成する。

[0045] 図1の例の正規化では、まず、第1正規化部4aが、時間周波数毎に、当該複素ベクトルが具備する特定の1つの要素を基準として当該複素ベクトルの各要素の偏角を正規化する。これにより、複素ベクトルの各要素の偏角は、源信号の位相や振幅に依存せず、各センサに対する信号源の相対位置や周波数のみに依存することになる（詳細は後述）。次に、第2正規化部4bが、第1正規化部4aで正規化された各要素の偏角を周波数に比例した値で除算する。これにより、各複素ベクトルの各要素の周波数依存性が排除され、複素ベクトルは、各センサに対する各信号源の相対位置のみに依存したものに正規化される。さらに、第3正規化部4cが、第2正規化部4bで正規化された各要素からなるベクトルのノルムを所定の値に正規化する。

[0046] 次に、クラスタリング部5が、このような正規化が行われた正規化ベクトルのクラスタリングを行い、クラスタを生成する。これらのクラスタは、各センサに対する各信号源の相対位置のみに依存したものになる。分離信号生成部6は、これらのクラスタを利用して各種の信号分離処理を行い、周波数領域の分離信号を生成する。最後に、時間領域変換部が、周波数領域の分離信号を時間領域の分離信号に変換する。

上述のように、これらのクラスタを生成するために、各センサの厳密な位置情報を事前に取得しておく必要はない。また、これらのクラスタの生成には、全てのセンサでの観測信号の情報が用いられている。すなわち、本発明では、複数のセンサによる観測信号から得られた情報を容易かつ効率的に利用して信号分離処理を行うことができる。

[0047] なお、ノルムを正規化しなくても、クラスタリングの方法を工夫することにより、各センサに対する各信号源の相対位置のみに依存したクラスタを生成することは可能である。しかし、クラスタリングの処理を簡略化するため、第3正規化部4cによるノルムの正規化を行うことが望ましい。

以下、本発明の各実施の形態を詳細に説明する。

〔第1の実施の形態（「第1の本発明」の例）〕

まず、本発明における第1の実施の形態について説明する。

[0048] 本形態は、上述の原理により、厳密なセンサ配置に係る情報の事前取得や煩雑な操作を必要とせず、正確にパーミュテーション問題を解決する形態である。なお、本

形態では、後述する「基底ベクトル」が上述の「複素ベクトル」に相当する。

＜ハードウェア構成＞

図2は、第1の実施の形態における信号分離装置10のハードウェア構成を例示したブロック図である。

図2に例示するように、この例の信号分離装置10は、CPU (Central Processing Unit) 10a、入力部10b、出力部10c、補助記憶装置10f、RAM (Random Access Memory) 10d、ROM (Read Only Memory) 10e及びバス10gを有している。

- [0049] この例のCPU10aは、制御部10aa、演算部10ab及びレジスタ10ac有し、レジスタ10acに読み込まれた各種プログラムに従って様々な演算処理を実行する。また、この例の入力部10bは、データが入力される入力ポート、キーボード、マウス等であり、出力部10cは、データを出力する出力ポート、ディスプレイ等である。補助記憶装置10fは、例えば、ハードディスク、MO (Magneto-Optical disc)、半導体メモリ等であり、本形態の信号分離処理を実行するための信号分離プログラムを格納した信号分離プログラム領域10fa及びセンサで観測された時間領域の混合信号等の各種データが格納されるデータ領域10fbを有している。また、RAM10dは、例えば、SRAM (Static Random Access Memory)、DRAM (Dynamic Random Access Memory) 等であり、信号分離プログラムが書き込まれる信号分離プログラム領域10da及び各種データが書き込まれるデータ領域10dbを有している。また、この例のバス10gは、CPU10a、入力部10b、出力部10c、補助記憶装置10f、RAM10d及びROM10eを通信可能に接続している。

[0050] ＜ハードウェアとソフトウェアとの協働＞

この例のCPU10aは、読み込まれたOS (Operating System) プログラムに従い、補助記憶装置10fの信号分離プログラム領域10faに格納されている信号分離プログラムを、RAM10dの信号分離プログラム領域10daに書き込む。同様にCPU10aは、補助記憶装置10fのデータ領域10fbに格納されている時間領域の混合信号等の各種データをRAM10dのデータ領域10dbに書き込む。さらに、CPU10aは、この信号分離プログラムや各種データが書き込まれたRAM10d上のアドレスをレジスタ10acに格納する。そして、CPU10aの制御部10aaは、レジスタ10acに格納されたこれ

らのアドレスを順次読み出し、読み出したアドレスが示すRAM10d上の領域からプログラムやデータを読み出し、そのプログラムが示す演算を演算部10abに順次実行させ、その演算結果をレジスタ10acに格納していく。

[0051] 図3は、このようにCPU10aに信号分離プログラムが読み込まれることにより構成される信号分離装置10のブロック図の例示である。また、図4Aは、図3におけるパーミュテーション問題解決部140の詳細を例示したブロック図であり、図4Bは、図4Aの基底ベクトル正規化部142の詳細を例示したブロック図である。

図3に例示するように、信号分離装置10は、メモリ100、周波数領域変換部120、分離行列算出部130、パーミュテーション問題解決部140、分離信号生成部150、時間領域変換部160及び制御部170を有している。また、この例のパーミュテーション問題解決部140は、逆行列算出部141(「複素ベクトル生成部」に対応)、基底ベクトル正規化部142(「正規化部」に対応)、クラスタリング部143、順列算出部144及び並び替え部145を有しており、基底ベクトル正規化部142は、周波数正規化部142a及びノルム正規化部142bを有している。また、周波数正規化部142aは、第1正規化部142aaと第2正規化部142abとを有している。さらに、制御部170は一時メモリ171を有している。

[0052] ここでメモリ100及び一時メモリ171は、レジスタ10ac、補助記憶装置10fのデータ領域10fb或いはRAM10dのデータ領域10db等に相当する。また、周波数領域変換部120、分離行列算出部130、パーミュテーション問題解決部140、分離信号生成部150、時間領域変換部160及び制御部170は、CPU10aにOSプログラムや信号分離プログラムが読み込まれることにより構成されるものである。

なお、図3及び図4における破線の矢印は理論上の情報の流れを示し、実線の矢印は実際のデータの流れを示す。また、これらの図において制御部170に出入りするデータの流れに対応する矢印は省略してあり、図4における実際のデータの流れに対応する矢印も省略してある。

[0053] <処理>

次に、本形態の信号分離装置10の処理について説明する。なお、以下では、N個の源信号が混合され、M個のセンサで観測された状況を取り扱う。また、前処理にお

いて、各センサで観測された時間領域の混合信号 $x_q(t)$ ($q=1, \dots, M$)がメモリ100の記憶領域101に格納され、信号の伝達速度 c 、 M 以下の自然数から選択された基準値 Q (M 個のセンサから選択された1つの基準センサの添字番号)及び実数 d の各パラメータが記憶領域107に格納されているものとする。

- [0054] 図5は、第1の実施の形態における信号分離装置10の処理の全体を説明するためのフローチャートである。以下、この図に沿って、本形態における信号分離装置10の処理を説明していく。

[周波数領域変換部120の処理]

まず、周波数領域変換部120において、メモリ100の記憶領域101から時間領域の混合信号 $x_q(t)$ を読み出し、これらを短時間離散フーリエ変換等によって周波数毎の時系列の信号(「周波数領域の混合信号」と呼ぶ) $X_q(f, \tau)$ ($q=1, \dots, M$)に変換し、メモリ100の記憶領域102に格納する(ステップS1)。

[0055] [分離行列算出部130の処理]

次に、分離行列算出部130において、メモリ100の記憶領域102から周波数領域の混合信号 $X_q(f, \tau)$ を読み出す。周波数領域の混合信号 $X_q(f, \tau)$ を読み出した分離行列算出部130は、これらからなる混合信号ベクトル $X(f, \tau)=[X_1(f, \tau), \dots, X_M(f, \tau)]^T$ を用い、独立成分分析(ICA)によって、周波数毎に第1の分離行列 $W(f)$ と分離信号ベクトル $Y(f, \tau)=[Y_1(f, \tau), \dots, Y_N(f, \tau)]^T$ とを算出する。そして、算出された第1の分離行列 $W(f)$ はメモリ100の記憶領域103に格納される(ステップS2)。

- [0056] ここで、分離行列算出部130において算出された第1の分離行列 $W(f)$ には、順序の任意性が含まれている。よって、以下のようにパーミュテーション問題解決部140において、第1の分離行列 $W(f)$ を順序の任意性を解決した第2の分離信号 $W'(f)$ に変更する。

[パーミュテーション問題解決部140の処理]

まず、逆行列算出部141において、メモリ100の記憶領域103から第1の分離行列 $W(f)$ を読み出し、そのムーア・ペンローズ(Moore-Penrose)型一般化逆行列 $W^+(f)=[A_1(f), \dots, A_N(f)]$ ($M=N$ の場合は逆行列 $W^{-1}(f)$ に一致)を計算し、これを構成する基底ベクトル $A_p(f)=[A_{1p}(f), \dots, A_{Mp}(f)]^T$ ($p=1, \dots, N$)をメモリ100の記憶領域104に格納する

(ステップS3)。なお、 $M=N$ の場合、 $W^+(f)$ は逆行列 $W^{-1}(f)$ に一致する。

[0057] 次に、基底ベクトル正規化部142において、メモリ100の記憶領域104から基底ベクトル $A_p(f)$ ($p=1, \dots, N, f=0, f/L, \dots, f_s(L-1)/L$)を読み出し、これらを正規化した正規化基底ベクトル $A_p''(f)$ を算出し、メモリ100の記憶領域106に格納する(ステップS4)。なお、基底ベクトル正規化部142は、全ての基底ベクトル $A_p(f)$ ($p=1, \dots, N, f=0, f/L, \dots, f_s(L-1)/L$)を正規化し、それらを周波数に依存せず、信号源の位置のみに依存する正規化基底ベクトル $A_p''(f)$ に正規化する。その結果、それらをクラスタリングした際に、各クラスタが各信号源に対応ようになる。この正規化を適切に行わないとクラスタが形成されない。また、本形態における正規化は周波数正規化とノルム正規化との二段階からなる。周波数正規化は、周波数正規化部142a(図4B)において、基底ベクトルを周波数に依存しない周波数正規化ベクトルに正規化するものである。ノルム正規化は、ノルム正規化部142bにおいて、周波数正規化ベクトルをノルムが規定値(この例では1)をとる正規化基底ベクトルに正規化するものである。これらの正規化処理の詳細については後述する。

[0058] 次に、クラスタリング部143において、メモリ100の記憶領域106から正規化基底ベクトル $A_p''(f)$ を読み出し、これらの正規化基底ベクトル $A_p''(f)$ をクラスタリングしてN個のクラスタ C_k ($k=1, \dots, N$)を生成し、各クラスタ C_k とそれらのセントロイド(中心ベクトル) η_k を特定する情報をメモリ100の記憶領域108, 109に格納する(ステップS5)。このクラスタリングは、例えば、各クラスタ C_k の要素(正規化基底ベクトル $A_v''(f)$)と各クラスタ C_k のセントロイド η_k との間の二乗和 U_k の総和 U

[0059] [数13]

$$U = \sum_{k=1}^M U_k$$

$$U_k = \sum_{A_v''(f) \in C_k} \|A_v''(f) - \eta_k\|^2$$

を最小値化することを基準に行われる。この最小値化は、例えば、非特許文献6などで解説されているk-meansクラスタリングを用いることによって効果的に行うことができる。なお、クラスタ C_k のセントロイド η_k は、

[0060] [数14]

$$\eta_k = \frac{\sum_{A_v''(f) \in C_k} A_v''(f) / |C_k|}{\left\| \sum_{A_v''(f) \in C_k} A_v''(f) / |C_k| \right\|}$$

によって計算される。ここで $|C_k|$ はクラスタ C_k における要素(正規化基底ベクトル $A_v''(f)$)の数である。また、ここでは、距離としてユークリッド距離の自乗を用いているが、これを一般化したミンコフスキー距離などを用いてもよい。なお、正規化基底ベクトル $A_p''(f)$ がクラスタを形成する理由については後述する。

次に、順列算出部144において、メモリ100の記憶領域106から正規化基底ベクトル $A_p''(f)$ を、記憶領域109から各クラスタ C_k のセントロイド η_k をそれぞれ読み出す。そして、順列算出部144は、これらを用い、第1の分離行列 $W(f)$ の要素を並び替えるための順列 Π_f ($\{1, 2, \dots, N\}$ から $\{1, 2, \dots, N\}$ への全単射な関数)を周波数 f 毎に算出し、メモリ100の記憶領域110に格納する(ステップS6)。なお、この順列 Π_f は、

[0061] [数15]

$$\Pi_f = \arg \min_{\Pi} \sum_{k=1}^N \left\| \eta_k - A_{\Pi(k)}''(f) \right\|^2 \quad \dots (13)$$

によって決定される。ここで式(13)における $\arg \min_{\Pi}$ は、 \cdot を最小値化する Π を意味する。また、この式における $A_{\Pi(k)}''(f)$ は、 Π によって正規化基底ベクトル $A_k''(f)$ に並び替えられる正規化基底ベクトルを意味する。すなわち、 $\Pi(k)$ 行目の正規化基底ベクトル $A_{\Pi(k)}''(f)$ は、 Π_f によって k 列目の正規化基底ベクトル $A_k''(f)$ となる。なお、式(13)に従って順列 Π_f を決定する手順としては、例えば、採り得るすべての順列 Π ($N!$ 通り)に対して

[0062] [数16]

$$\sum_{k=1}^N \left\| \eta_k - A_{\Pi(k)}''(f) \right\|^2$$

を算出し、その最小値に対応する Π を順列 Π_f として決定する手順を例示できる。以下にこの具体例を示す。

[順列 Π_f 決定の具体例1]

信号源の数 N が3であり、ある周波数 f における正規化基底ベクトル $A_1''(f)$, $A_2''(f)$, $A_3''(f)$ と各セントロイド η_1 , η_2 , η_3 との距離の自乗が、以下の表に示されるもので

あったとする。

[0063] [表1]

	$A_1''(f)$	$A_2''(f)$	$A_3''(f)$
η_1	0.85	0.1	0.7
η_2	0.9	0.6	0.2
η_3	0.15	0.8	0.95

この場合、式(13)により得られる順列は、

$$\Pi_f: [1,2,3] \rightarrow [2,3,1]$$

となる。なぜなら、

$$\begin{aligned} \|\eta_1 - A_{\pi(1)}''(f)\|^2 &= \|\eta_1 - A_2''(f)\|^2 = 0.1 \\ \|\eta_2 - A_{\pi(2)}''(f)\|^2 &= \|\eta_2 - A_3''(f)\|^2 = 0.2 \\ \|\eta_3 - A_{\pi(3)}''(f)\|^2 &= \|\eta_3 - A_1''(f)\|^2 = 0.15 \end{aligned}$$

という組合せが、

[0064] [数17]

$$\sum_{k=1}^3 \|\eta_k - A_{\Pi(k)}''(f)\|^2$$

を最小とするからである(順列 Π_f 決定の具体例1の説明終わり)。

しかし、この手順はNが大きくなると現実的ではない。そのため、その近似法として、順番に $\|\eta_k - A_{\Pi(k)}''(f)\|^2$ を最小値化する $A_{\Pi(k)}''(f)$ を重複がないように選択していき、この選択された $A_{\Pi(k)}''(f)$ を正規化基底ベクトル $A_k''(f)$ に移す順列を順列 Π_f とする手順等を用いてもよい。以下、上述の「順列 Π_f 決定の具体例1」と同じ条件に対し、この近似法を適用して順列 Π_f を決定する手順を示す。

[0065] [順列 Π_f 決定の具体例2]

まず、上記の表1の場合、距離の自乗の最小値は0.1(正規化基底ベクトル $A_2''(f)$ と各セントロイド η_1 との距離の自乗)であるため、 $\Pi(1)=2$ を決定する。そして、正規化基底ベクトル $A_2''(f)$ と各セントロイド η_1 に関連する行と列とを消すと以下のようになる。

[0066] [表2]

	$A_1''(f)$	$A_2''(f)$	$A_3''(f)$
η_1			
η_2	0.9		0.2
η_3	0.15		0.95

この表2の場合、距離の自乗の最小値は0.15(正規化基底ベクトル $A_1''(f)$ と各セントロイド η_3 との距離の自乗)であるため、 $\Pi(3)=1$ を決定する。そして、最後に $\Pi(2)$ に残りの3を割り当てる([順列 Π_f 決定の具体例2]の説明終わり)。

次に、並び替え部145において、メモリ100の記憶領域103から第1の分離行列 $W(f)$ を、記憶領域110から順列 Π_f を読み出す。そして、並び替え部145は、この第1の分離行列 $W(f)$ の行を順列 Π_f に従って並び替えた第2の分離行列 $W'(f)$ を生成してメモリ100の記憶領域111に格納する(ステップS7)。なお、第1の分離行列 $W(f)$ の行を順列 Π_f に従って並び替えるとは、上述のムーア・ペンローズ型一般化逆行列 $W^+(f)$ における要素 $A_{\Pi(k)}''(f)$ から $A_k''(f)$ への並び替えに対応する並び替えを、第1の分離行列 $W(f)$ に対して行うことを意味する。すなわち、第1の分離行列 $W(f)$ の順列 $\Pi_f(k)$ 行目が第2の分離行列 $W'(f)$ の k 行目になるように並び替える。上述の[順列 Π_f 決定の具体例1, 2]の場合、第1の分離行列 $W(f)$ の2, 3, 1行目が、それぞれ第2の分離行列 $W'(f)$ の1, 2, 3行目になる。

[0067] [分離信号生成部150の処理]

その後、分離信号生成部150において、メモリ100の記憶領域102から周波数領域の混合信号 $X_q(f, \tau)$ を、記憶領域111から第2の分離行列 $W'(f)$ を読み出す。そして、分離信号生成部150は、周波数領域の混合信号 $X_q(f, \tau)$ からなる混合信号ベクトル $X(f, \tau)=[X_1(f, \tau), \dots, X_M(f, \tau)]^T$ と第2の分離行列 $W'(f)$ とを用い、分離信号ベクトル

$$Y(f, \tau) = W'(f) \cdot X(f, \tau)$$

を算出し、これらの各要素である周波数領域の信号(「周波数領域の混合信号」と呼ぶ) $Y_p(f, \tau)$ をメモリ100の記憶領域112に格納する(ステップS8)。

[0068] [時間領域変換部160の処理]

最後に時間領域変換部160において、メモリ100の記憶領域112から周波数領域

の分離信号 $Y_p(f, \tau)$ を読み出し、これを添え字 p 毎に($Y_p(f, \tau)$ 毎に)短時間逆フーリエ変換等によって時間領域の分離信号 $y_p(t)$ に変換し、これらの時間領域の分離信号 $y_p(t)$ をメモリ100の記憶領域113に格納する(ステップS9)。

[正規化処理の詳細(ステップS4の詳細)]

次に、前述した基底ベクトル正規化部142における正規化処理の詳細(ステップS4の詳細)について説明する。

[0069] 図6は、この正規化処理の詳細を説明するためのフローチャートである。

まず、制御部170(図3)においてパラメータ p に1を代入し、これを一時メモリ171に格納する(ステップS11)。また、制御部170においてパラメータ q に1を代入し、これを一時メモリ171に格納する(ステップS12)。次に、周波数正規化部142a(図4)において、メモリ100の記憶領域107から前述のパラメータ d, c, Q を読み出し、一時メモリ171からパラメータ p, q を読み出し、基底ベクトル $A_p(f)$ の要素 $A_{qp}(f)$ に対し、

[0070] [数18]

$$A_{qp}'(f) = |A_{qp}(f)| \exp \left[j \frac{\arg[A_{qp}(f)/A_{Qp}(f)]}{4fc^{-1}d} \right] \quad \dots (14)$$

の演算を行い、当該演算結果 $A_{qp}'(f)$ を周波数正規化ベクトル $A_p'(f)$ の各要素 $A_{qp}'(f)$ としてメモリ100の記憶領域105に格納する(ステップS13)。なお、 $\arg[\cdot]$ は偏角を意味し、 j は虚数単位を意味する。

より詳細には、まず、周波数正規化部142aの第1正規化部142aaが、基底ベクトル $A_p(f)$ の特定の1つの要素 $A_{Qp}(f)$ を基準として当該基底ベクトル $A_p(f)$ の各要素 $A_{qp}(f)$ の偏角を、以下の演算によって正規化する。

[0071] [数19]

$$A_{qp}'''(f) = |A_{qp}(f)| \exp \{ j \cdot \arg[A_{qp}(f)/A_{Qp}(f)] \} \quad \dots (15)$$

次に、周波数正規化部142aの第2正規化部142abが、第1正規化部142aaで正規化された各要素 $A_{qp}'''(f)$ の偏角を、以下のように周波数 f に比例した値 $4fc^{-1}d$ で除算する。

[0072] [数20]

$$A_{qp}'(f) = |A_{qp}'''(f)| \exp \left[j \frac{\arg[A_{qp}'''(f)]}{4fc^{-1}d} \right] \quad \dots (16)$$

次に、制御部170において、一時メモリ171に格納されたパラメータ q が $q=M$ を満たすか否かを判断する(ステップS14)。ここで $q=M$ でなければ、制御部170は、 $q+1$ の演算結果を新たなパラメータ q の値とし、これを一時メモリ171に格納し(ステップS15)、処理をステップS13へ戻す。一方、 $q=M$ であれば、制御部170は、さらに $p=N$ を満たすか否かを判断する(ステップS16)。

[0073] ここで $p=N$ でなければ、制御部170において、 $p+1$ の演算結果を新たなパラメータ p の値とし、これを一時メモリ171に格納し(ステップS17)、処理をステップS12へ戻す。一方、 $p=N$ であれば、制御部170においてパラメータ p に1を代入し、これを一時メモリ171に格納し(ステップS18)、ノルム正規化部142bにおける処理を開始する。まず、ノルム正規化部142bにおいて、メモリ100の記憶領域105から周波数正規化ベクトル $A_p'(f)$ の各要素 $A_{pq}'(f)$ を読み出し、

[0074] [数21]

$$\|A_p'(f)\| = \sqrt{\sum_{q=1}^M (A_{pq}'(f))^2} \quad \dots (17)$$

の演算を行って、周波数正規化ベクトル $A_p'(f)$ のノルム $\|A_p'(f)\|$ を求め、周波数正規化ベクトル $A_p'(f)$ 及びそのノルム $\|A_p'(f)\|$ を一時メモリ171に格納する(ステップS19)。

次に、ノルム正規化部142bにおいて、一時メモリ171から周波数正規化ベクトル $A_p'(f)$ 及びそのノルム $\|A_p'(f)\|$ を読み出し、

$$A_p''(f) = A_p'(f) / \|A_p'(f)\| \quad \dots (18)$$

の演算を行って正規化基底ベクトル $A_p''(f)$ を求め、これをメモリ100の記憶領域106に格納する(ステップS20)。

[0075] その後、制御部170において、一時メモリ171に格納されたパラメータ p が $p=N$ を満たすか否かを判断する(ステップS21)。ここで $p=N$ でなければ、制御部170は、 $p+1$ の演算結果を新たなパラメータ p の値とし、これを一時メモリ171に格納し(ステッ

プS22)、ステップS19の処理に戻す。一方、 $p=N$ であれば、制御部170は、ステップS4の処理を終了させる。

このように生成された正規化基底ベクトル $A_p''(f)$ は、周波数に依存せず、信号源の位置のみに依存するベクトルとなる。その結果、この正規化基底ベクトル $A_p''(f)$ はクラスタを形成することになる。以下にこの理由を説明する。

[0076] [正規化基底ベクトル $A_p''(f)$ がクラスタを形成する理由]

基底ベクトル $A_p(f)$ の各要素 $A_{qp}(f)$ は、源信号 p に対応する信号源 k からセンサ q への周波数応答 H_{qk} に比例した(ある複素数スカラーが掛かった)ものになっている。これらの複素数スカラーは離散時間に応じて(すなわち位相に応じて)変化するが、源信号 p とセンサ q とに対応する複素数スカラーと、源信号 p とセンサ Q とに対応する複素数スカラーとの相対値は、離散時間が変化しても一定である(周波数 f が同じであれば)。すなわち、周波数 f が同じであれば、源信号 p とセンサ q とに対応する複素数スカラーの偏角と、源信号 p とセンサ Q とに対応する複素数スカラーの偏角との相対値は一定である。

[0077] 前述のように、周波数正規化部142aの第1正規化部142aaは、基底ベクトル $A_p(f)$ の特定の1つの要素 $A_{qp}(f)$ を基準として当該基底ベクトル $A_p(f)$ の各要素 $A_{qp}(f)$ の偏角を正規化する。これにより、上述の複素数スカラーの位相に起因する不確定性を取り除き、源信号 p とセンサ q とに対応する基底ベクトル $A_p(f)$ の要素 $A_{qp}(f)$ の偏角を、源信号 p と基準センサ Q (基準値 Q に対応)とに対応する基底ベクトル $A_p(f)$ の要素 $A_{qp}(f)$ の偏角に対する相対値として表現する。なお、この場合、要素 $A_{qp}(f)$ の偏角に対応する当該相対値は0と表現される。信号源 k からセンサ q への周波数応答を、反射や残響の無い直接波モデルで近似して考える。すると、上記の第1正規化部142aaにより正規化された偏角は、信号源 k からセンサへの波の到達時間差と周波数 f の双方に比例したものになる。ここでの到達時間差とは、信号源 k からの波がセンサ q に到達するまでの時間と、当該波が基準センサ Q に到達するまでの時間との時間差である。

[0078] また、前述のように、第2正規化部142abは、第1正規化部142aaで正規化された各要素 $A_{qp}''(f)$ の偏角を周波数 f に比例した値で除算する。これにより、各要素 $A_{qp}''(f)$ を、それらの偏角の周波数依存性を排除した各要素 $A_{qp}'(f)$ に正規化する。これにより、正規化された各要素 $A_{qp}'(f)$ は、直接波モデルに従えば、信号源 k からセンサ

への波の到達時間差のみに依存するものとなる。ここで、信号源 k からセンサへの波の到達時間差は、信号源 k 、センサ q 、基準センサ Q の相対位置にのみ依存する。そのため、信号源 k 、センサ q 、基準センサ Q が同じであれば、周波数 f が異なっても各要素 $A'_{qp}(f)$ の偏角は同一となる。従って、周波数正規化ベクトル $A'_p(f)$ は、周波数 f には依存せず、信号源 k の位置のみに依存する。

[0079] そのため、周波数正規化ベクトル $A'_p(f)$ のノルムを正規化した正規化基底ベクトル $A''_p(f)$ のクラスタリングによって、同じ信号源毎に対応するクラスタが形成される。なお、実際の環境では、反射や残響などの影響により、直接波モデルは厳密には満たされないが、後述の実験結果に示すとおり十分に良い近似となっている。

次に、正規化基底ベクトル $A''_p(f)$ がクラスタを形成する理由を、モデルを用いて説明する。前述した式(1)に示されるインパルス応答 $h_{qk}(r)$ を直接波(近距離場)混合モデルを用いて近似し、周波数領域で表すと

[0080] [数22]

$$H_{qk}(f) = \frac{1}{d_{qk}} \exp[-j2\pi f c^{-1}(d_{qk} - d_{Qk})] \quad \dots (19)$$

となる。ここで、 d_{qk} は信号源 k とセンサ q との間の距離であり、また、減衰 $1/d_{qk}$ は、距離 d_{qk} によって決まり、遅延 $(d_{qk} - d_{Qk})/c$ は、基準センサ Q の位置で正規化された距離によって決まる。

また、独立成分分析(ICA)における順序の任意性とスケーリングの任意性を考慮すると、基底ベクトル $A_p(f)$ と式(4)における信号源 k から各センサへの周波数応答をまとめたベクトル $H_k(f)$ の間には、以下の関係が成り立つ。

[0081] $A_p(f) = \varepsilon_p \cdot H_k(f), \quad A_{qp}(f) = \varepsilon_p \cdot H_{qk}(f) \quad \dots (20)$

ここで、 ε_p はスケーリングの任意性を表現する複素スカラー値であり、添え字 k と q が異なる可能性が順序の任意性を表現している。ここで、式(16)(18)(19)(20)より、

[0082] [数23]

$$A_{qp}''(f) = \frac{1}{d_{qk} D} \exp\left[-j \frac{\pi (d_{qk} - d_{Qk})}{2d}\right], \quad D = \sqrt{\sum_{i=1}^M \frac{1}{d_{ik}^2}} \quad \dots (21)$$

となる。この式から分かるように、正規化基底ベクトル $A_p''(f)$ の各要素 $A_{qp}''(f)$ は、周波数 f からは独立であり、信号源 k とセンサ q の位置のみに依存する。従って、正規化基底ベクトル $A_p''(f)$ をクラスタリングすると、同じ信号源ごとに対応するクラスが形成される。

同様なことは、信号の減衰を考慮していない近距離場混合モデルでモデル化した場合にも言える。すなわち、前述した式(1)に示される畳み込み混合モデルを、減衰を無視した近距離場混合モデルで近似し、周波数領域で表すと

[0083] [数24]

$$H_{qk}(f) = \exp[-j2\pi f c^{-1}(d_{qk} - d_{Qk})] \quad \dots (22)$$

となる。この場合、式(16)(18)(20)(22)より、

[0084] [数25]

$$A_{qp}''(f) = \frac{1}{\sqrt{M}} \exp\left[-j\frac{\pi}{2} \frac{(d_{qk} - d_{Qk})}{d}\right] \quad \dots (23)$$

となる。この場合も、正規化基底ベクトル $A_p''(f)$ の各要素 $A_{qp}''(f)$ は、周波数 f からは独立であり、信号源 k とセンサ q の位置のみに依存している。

さらに、近距離場混合モデルだけではなく、遠距離場混合モデルでも同様なことがいえる。すなわち、前述した式(1)に示される畳み込み混合モデルを遠距離場混合モデルで近似し、周波数領域で表すと

[0085] [数26]

$$H_{qk}(f) = \exp[-j2\pi f c^{-1} \|SE_q - SE_Q\| \cos\theta_k^{qQ}] \quad \dots (24)$$

となる。なお、 SE_q 及び SE_Q はセンサ q , Q の位置を示すベクトルである。また θ_k^{qQ} は、センサ q , Q を結ぶ直線と、センサ q , Q の中心点と信号源 k とを結ぶ直線と、で成す角度を示す。この場合、式(16)(18)(20)(24)より、

[0086] [数27]

$$A_{qp}''(f) = \frac{1}{\sqrt{M}} \exp\left[-j\frac{\pi}{2} \frac{\|SE_q - SE_Q\| \cos\theta_k^{qQ}}{d}\right] \quad \dots (25)$$

となる。この場合も、正規化基底ベクトル $A_p''(f)$ の各要素 $A_{qp}''(f)$ は、周波数 f からは独立であり、信号源 k とセンサ q の位置のみに依存している。

また、式(21)からパラメータ d の値は、 $d > d_{\max}/2$ であることが望ましく(d_{\max} は要素 $A_{qp}''(f)$ に対応する基準センサ Q と他のセンサとの最大距離 d_{\max} を意味する。)、より好ましくは $d \geq d_{\max}$ であることが望ましく、さらにより好ましくは $d = d_{\max}$ であることが望ましいことが分かる。以下、この理由について説明する。

[0087] 図7及び図8は、パラメータ d 毎の正規化基底ベクトルの要素 $A_{qp}''(f)$ とその偏角 $\arg[A_{qp}''(f)]$ との関係を説明するための複素平面図である。なお、これらにおける横軸は実軸を縦軸は虚軸を示している。

図7Aは $d_{\max}/2 \geq d$ の場合における複素平面図である。ここで上述の d_{\max} の定義より、任意の q 及び k に対し $d_{qk} - d_{Qk}$ の絶対値は d_{\max} 以下となる。よって $d_{\max}/2 \geq d$ の場合、 $(\pi/2) \cdot (d_{qk} - d_{Qk})/d \leq -\pi$ 、 $(\pi/2) \cdot (d_{qk} - d_{Qk})/d \geq \pi$ と成り得る。その結果、式(21)で表される $A_{qp}''(f)$ の偏角 $\arg[A_{qp}''(f)]$ は、 2π を超える $\alpha_1 \leq \arg[A_{qp}''(f)] \leq \alpha_2$ ($\alpha_1 \leq -\pi$ 、 $\alpha_2 \geq \pi$)の範囲に分布する可能性がある。そのため、異なる正規化基底ベクトルの要素 $A_{qp}''(f)$ の偏角が一致する可能性があり、前述したクラスタリングにおいて、異なる正規化基底ベクトル $A_p''(f)$ を同じクラスにクラスタリングしてしまう可能性がある。よって、 $d > d_{\max}/2$ であることが望ましい。しかし、この偏角の重複範囲に対応する正規化基底ベクトル $A_p''(f)$ のサンプルが存在しないならば $d_{\max}/2 \geq d$ としても問題はない。

[0088] 図7Bは $d_{\max}/2 < d < d_{\max}$ の場合における複素平面図である。この場合、 $-\pi < (\pi/2) \cdot (d_{qk} - d_{Qk})/d < -\pi/2$ 、 $\pi/2 < (\pi/2) \cdot (d_{qk} - d_{Qk})/d < \pi$ と成り得る。その結果、式(21)で表される $A_{qp}''(f)$ の偏角 $\arg[A_{qp}''(f)]$ は、 $\beta_1 \leq \arg[A_{qp}''(f)] \leq \beta_2$ ($-\pi < \beta_1 < -\pi/2$ 、 $\pi/2 < \beta_2 < \pi$)の範囲に分布する可能性がある。そのため、 $-\pi < \arg[A_{qp}''(f)] < -\pi/2$ 及び $\pi/2 < \arg[A_{qp}''(f)] < \pi$ の範囲において、異なる周波数正規化ベクトル要素間の偏角の差の増加に伴い、これらの周波数正規化ベクトル要素間の距離が単調増加しないこともありうる。これは、前述したクラスタリングの精度を低下させる可能性がある。よって $d \geq d_{\max}$ であることがより望ましい。

[0089] 図8Aは $d = d_{\max}$ の場合における複素平面図であり、図8Bは $d > d_{\max}$ の場合における

複素平面図である。ここで $d > d_{\max}$ の場合、 $-\pi/2 < (\pi/2) \cdot (d - d_{\max})/d < 0$, $0 < (\pi/2) \cdot (d - d_{\max})/d < \pi/2$ と成り得る。その結果、式(21)で表される $A''(f)$ の偏角 $\arg[A''(f)]$ は、図8Bに示すように、 $\gamma_1 \leq \arg[A''(f)] \leq \gamma_2$ ($-\pi/2 < \gamma_1 < 0$, $0 < \gamma_2 < \pi/2$)の範囲に分布する。そして、 d が大きくなればなるほどその分布範囲は狭くなっていき、狭い範囲にクラスタが密集していく。これは、前述したクラスタリングの精度を低下させる。

- [0090] これに対し、 $d = d_{\max}$ である場合、 $-\pi/2 \leq (\pi/2) \cdot (d - d_{\max})/d < 0$, $0 < (\pi/2) \cdot (d - d_{\max})/d \leq \pi/2$ と成り得る。その結果、式(21)で表される $A''(f)$ の偏角 $\arg[A''(f)]$ は図8Aに示すように $-\pi/2 \leq \arg[A''(f)] \leq \pi/2$ の範囲に分布する。この場合、周波数正規化ベクトルの要素間における偏角の差の増加に対して、それらの距離も単調増加するという関係を維持しつつ、できるだけ広い範囲にクラスタを分散させることができる。その結果、一般的にクラスタリングの精度を向上させることができる。

〔第2の実施の形態(「第1の本発明」の例)〕

次に本発明における第2の実施の形態について説明する。

- [0091] 第1の実施の形態では、基底ベクトルから得られる情報によりパーミュテーション問題を解決したが、本形態では、これに特開2004-145172号公報や「H. Sawada, R. Mukai, S. Araki, S. Makino, "A Robust and Precise Method for Solving the Permutation Problem of Frequency-Domain Blind Source Separation," IEEE Trans. Speech and Audio processing, vol. 12, no. 5, pp. 530-538, Sep. 2004. (以下「参考文献」と呼ぶ)」で示されているような分離信号のエンベロープの情報を統合することで、より精度よくパーミュテーション問題を解決する。なお、これらの文献では、基底ベクトルの代わりに信号源の方向に関する情報が利用されている。

- [0092] 以下では、第1の実施の形態との相違点を中心に説明を行い、これと共通する事項については説明を省略する。

<構成>

図9は、第2の実施の形態における信号分離装置200のブロック図の例示である。なお、本形態の信号分離装置200もCPU10a(図2)に信号分離プログラムが読み込まれることにより構成されるものである。また、図10Aは、図9におけるパーミュテ-

ション問題解決部240の詳細を例示したブロック図であり、図10Bは、図10Aの順列修正部247の詳細を例示したブロック図である。なお、図9及び図10において第1の実施の形態と共通する部分については第1の実施の形態と同じ符号を付した。また、図9及び図10における破線の矢印は理論上の情報の流れを示し、実線の矢印は実際のデータの流れを示す。また、これらの図において制御部170に出入りするデータの流れに対応する矢印は省略してあり、図10における実際のデータの流れに対応する矢印も省略してある。

[0093] 本形態の第1の実施の形態のとの相違点は、主にパーミュテーション問題解決部240の構成である。すなわち、本形態のパーミュテーション問題解決部240は、第1の実施の形態のパーミュテーション問題解決部140に、順列評価部246と順列修正部247とを付加したものとなっている(図9、図10A)。なお、順列評価部246は、周波数毎に順列の信頼度を評価するものであり、順列修正部247は、順列の信頼度が低いと評価された周波数に対し、分離信号のエンベロープを用いて新たな順列を算出するものである。また、順列修正部247は、判定部247aと分離信号生成部247bとエンベロープ算出部247cと順列再算出部247dと再並び替え部247eを有している(図10B)。また、本形態では、順列算出部144と順列修正部247とによって請求項4の「順列算出部」を構成している。

[0094] <処理>

図11は、第2の実施の形態における信号分離装置200の処理の全体を説明するためのフローチャートである。以下、この図に沿って、本形態における信号分離装置200の処理を説明していく。

ステップS51からS57の処理は、第1の実施の形態におけるステップS1からS7までと同じであるため説明を省略する。本形態では、このステップS57の処理の後に、順列評価部246において、周波数毎に順列 Π_f の信頼度を評価し、順列 Π_f の信頼度が低いと評価された周波数に対し、分離信号のエンベロープを用いて順列 Π'_f を算出し、この周波数のみについて第2の分離行列 $W'(f)$ の行を順列 Π'_f に従って並び替えて第3の分離行列 $W''(f)$ を生成し、メモリ100の記憶領域110に格納する(ステップS58)。なお、この処理の詳細については後述する。

[0095] その後、分離信号生成部150において、メモリ100の記憶領域102から周波数領域の混合信号 $X_q(f, \tau)$ を、記憶領域111から第3の分離行列 $W''(f)$ を読み出す。そして、分離信号生成部150は、周波数領域の混合信号 $X_q(f, \tau)$ からなる混合信号ベクトル $X(f, \tau) = [X_1(f, \tau), \dots, X_M(f, \tau)]^T$ と第3の分離行列 $W''(f)$ とを用い、分離信号ベクトル

$$Y(f, \tau) = W''(f) \cdot X(f, \tau)$$

を算出し、周波数領域の分離信号 $Y_p(f, \tau)$ をメモリ100の記憶領域112に格納する(ステップS59)。

[0096] 最後に時間領域変換部160において、メモリ100の記憶領域112から周波数領域の分離信号 $Y_p(f, \tau)$ を読み出し、これを添え字 p 毎に時間領域の分離信号 $y_p(t)$ に変換し、これらの時間領域の分離信号 $y_p(t)$ をメモリ100の記憶領域113に格納する(ステップS60)。

[ステップS58の処理の詳細]

図12及び図13は、図11におけるステップS58の例を説明するためのフローチャートである。以下、このフローチャートに沿ってステップS58の詳細を説明していく。

[0097] まず、制御部170においてパラメータ f に0を代入し、集合 F を空集合とし、これらを示す情報を一時メモリ171に格納する(ステップS71)。次に、順列評価部246において、メモリ100の記憶領域110に格納された順列 Π_f の信頼度を周波数毎に評価し、その評価結果 $\text{trust}(f)$ を一時メモリ171に格納する(ステップS72)。ここで、順列 Π_f の信頼性が高いとは、例えば、正規化基底ベクトル $A_p''(f)$ が、それぞれに対応するセントロイド η_k に十分近いということである。また、正規化基底ベクトル $A_p''(f)$ が、それぞれに対応するセントロイド η_k に十分近いかどうかは、例えば、正規化基底ベクトル $A_p''(f)$ とセントロイド η_k との距離がクラス C_k の分散 $U_k / |C_k|$ より小さいかどうか、すなわち、

$$U_k / |C_k| > \| \eta_k - A_{\Pi(f)}''(f) \|^2 \quad \dots (26)$$

により判定できる。よって、このステップでは、例えばまず、順列評価部246において、メモリ100の記憶領域105から正規化基底ベクトル $A_p''(f)$ を、記憶領域109からセントロイド η_k を、記憶領域110から順列 Π_f をそれぞれ読み出す。そして、順列評価

部246は、周波数 f 毎に、式(26)を満たすか否かを判断し、満たす場合には $\text{trust}(f)=1$ を、満たさない場合には $\text{trust}(f)=0$ を出力し、一時メモリ171に格納する。

[0098] 次に、順列修正部247の判定部247aが、一時メモリ171から周波数 f 毎に評価結果 $\text{trust}(f)$ を読み出し、 $\text{trust}(f)=1$ であるか否かを判断する(ステップS73)。ここで、 $\text{trust}(f)=0$ である場合、そのままステップS76の処理に進む。一方、 $\text{trust}(f)=1$ である場合、制御部170において、集合 F と $\{f\}$ の和集合を新たな集合 F として一時メモリ171に格納し(ステップS74)、再並び替え部247eにおいて、この周波数 f における第2の分離行列 $W'(f)$ を第3の分離行列 $W''(f)$ としてメモリ100の記憶領域111に格納し(ステップS75)、ステップS76に進む。

[0099] ステップS76では、制御部170において、一時メモリ171に格納されたパラメータ f の値が $f=(L-1)f_s/L$ を満たすか否かを判断し(ステップS76)、これを満たさなければ f_s/L の演算結果を新たなパラメータ f の値として(ステップS77)一時メモリ171に格納し、ステップS72に戻る。

一方、 $f=(L-1)f_s/L$ を満たす場合、分離信号生成部247bにおいて、集合 F に属しない周波数 f を1つ選択する。そして、分離信号生成部247bは、この周波数 f とその近傍であって集合 F に属する全ての周波数 $g(g \in F, |g-f| \leq \delta, \delta$ は定数)とに対し、メモリ100の記憶領域102から周波数領域の混合信号 $X(f, \tau)=[X_1(f, \tau), \dots, X_M(f, \tau)]^T$, $X(g, \tau)=[X_1(g, \tau), \dots, X_M(g, \tau)]^T$ を読み出し、記憶領域111から第2の分離行列 $W'(f), W'(g)$ を読み出し、

$$Y(f, \tau)=W'(f) \cdot X(f, \tau)$$

$$Y(g, \tau)=W'(g) \cdot X(g, \tau)$$

によって分離信号 $Y(f, \tau)=[Y_1(f, \tau), \dots, Y_N(f, \tau)]^T$, $Y(g, \tau)=[Y_1(g, \tau), \dots, Y_N(g, \tau)]^T$ を算出して一時メモリ171に格納する(ステップS78)。

[0100] 次に、エンベロープ算出部247cにおいて、一時メモリ171から全ての周波数領域の分離信号 $Y_p(f, \tau), Y_p(g, \tau)$ を読み出し、これらのエンベロープ

$$v_p^f(\tau)=|Y_p(f, \tau)|$$

$$v_p^g(\tau)=|Y_p(g, \tau)|$$

を算出して一時メモリ171に格納する(ステップS79)。

そして、順列再算出部247dにおいて、周波数の差 δ 以下の近傍における相関corの和の最大値

[0101] [数28]

$$R_f = \max_{\Pi} \sum_{|g-f| \leq \delta} \sum_{k=1}^N \text{cor}(v_{\Pi(k)}^f, v_{\Pi'(k)}^g)$$

を算出して一時メモリに格納する(ステップS80)。ここで、 Π' は周波数 g に対し、既に決定された順列である。なお、この式における相関 $\text{cor}(\Phi, \Psi)$ は、2つの信号 Φ 、 Ψ の相関を示し、以下のように定義される。

$$\text{cor}(\Phi, \Psi) = (\langle \Phi, \Psi \rangle - \langle \Phi \rangle \cdot \langle \Psi \rangle) / (\sigma_{\Phi} \cdot \sigma_{\Psi})$$

ただし、 $\langle \cdot \rangle$ は \cdot の時間平均であり、 σ_{Φ} は Φ の標準偏差である。また、 $v_{\Pi(k)}^f$ は、 Π によってエンベロープ $v_k^f(\tau)$ に並び替えられるエンベロープを意味する。すなわち、 $\Pi(k)$ 列目のエンベロープ $v_{\Pi(k)}^f$ は、 Π' によって k 番目のエンベロープ $v_k^f(\tau)$ となる。

[0102] また、順列再算出部247dにおいて、この相関corの和を最大化する順列

[0103] [数29]

$$\Pi_f' = \arg \max_{\Pi} \sum_{|g-f| \leq \delta} \sum_{k=1}^N \text{cor}(v_{\Pi(k)}^f, v_{\Pi'(k)}^g)$$

を算出してメモリ100の記憶領域110に格納する(ステップS81)。ここで、 Π' は周波数 g に対し、既に決定された順列である。なお、 $\arg \max_{\Pi} v$ は、 v を最大化する Π を意味する。

次に、制御部170において集合 F と $\{\zeta\}$ ($\zeta = \arg \max_f R_f$)との和集合を新たな集合 F として一時メモリ171に格納する(ステップS82)。そして、再並び替え部247eにおいて、 $f = \zeta$ とし、第2の分離行列 $W'(\zeta)$ の行を順列 Π_f' に従って並び替えた第3の分離行列 $W''(\zeta)$ を生成し、これをメモリ100の記憶領域111に格納する(ステップS83)。

[0104] 次に、制御部170において、一時メモリ171に格納された集合 F が全ての離散化された周波数 $f=0, f/L, \dots, f(L-1)/L$ の要素を有するか否かを判断する(ステップS84)。ここで、集合 F が全ての離散化された周波数 $f=0, f/L, \dots, f(L-1)/L$ の要素を有しない

のであれば、制御部170は、処理をステップS78に戻す。一方、集合Fが全ての離散化された周波数 $f=0, f/L, \dots, f(L-1)/L$ の要素を有するのであれば、制御部170は、ステップS58の処理を終了させる。

なお、以上の方法を用いず、前述した特開2004-145172号公報や「参考文献」に記載されたその他の方法を用い、ステップS58の処理を実行してもよい。

[0105] <実験結果>

次に、上述した第1の実施の形態及び第2の実施の形態による音源分離実験結果を示す。

[実験結果1]

1つ目の実験は不規則なセンサ配置によるものである。実験条件は図14Aに示す通りである。ここでは4つの無指向性マイクを不規則に配置したものをを用いたが、これらの配置情報としては、マイク間隔の上限が4cmであるということだけを信号分離装置に与えた。音源の数は3つであり、3秒間の英語の音声をスピーカから流した。結果を図14BにSIR (signal-to-interference ratio) によって示す。大きいSIRほど分離性能が良いことを示す。ここではパーミュテーション問題を解決する4種類の手法に関して結果を比較した。Envは分離信号のエンベロープ $|Y_p(f, \tau)|$ の情報のみを用いたもの、Basisは正規化基底ベクトル $A_p''(f)$ のクラスタリングによるもの(第1の実施の形態の手法)、Basis+Envはそれら2種類の情報を統合してより精度良く問題を解決したものの(第2の実施の形態の手法)、Optimalは源信号 s_k とインパルス応答 $h_{ak}(r)$ を知って得た最適な順列によるものである。

[0106] 結果を考察すると、Envだけでは性能にばらつきがあるが、第1の実施の形態のBasisはそれだけで十分な分離性能を得ている。また、双方を統合した第2の実施の形態のBasis+Envの結果はほぼOptimalに近い。このように不規則なセンサ配置を用いた場合でも、本発明により、周波数領域でのブラインド信号分離が高性能で達成された。

[実験結果2]

2つ目の実験は規則的なセンサ配置によるものである。実験条件は図15Aに示す通りである。3つの無指向性マイクを直線状に4cm間隔で配置した。音源の数は同

様に3つであり、3秒間の英語の音声をスピーカから流した。結果を図15Bに示す。この実験では、従来技術として示した信号源の位置に関する推定値をクラスタリングする方法も比較対象とし、6種類の手法に関して結果を比較した。DOAは信号の到来方向(DOA:direction-of-arrival)の推定値のみでパーミュテーション問題を解決したもの、DOA+Envはさらに分離信号のエンベロープ情報を統合したものがある。

- [0107] 従来技術であるDOA及びDOA+Envと、本発明であるBasis及びBasis+Envの結果を比較すると、従来手法が適用可能である規則的なセンサ配置に対しても、総じて本発明により結果が改善されていることがわかる。なお、演算量に関しては、従来技術とほぼ同等であった。

<第1, 2の実施の形態の特徴>

以上より、上述した第1, 2の実施の形態の特徴をまとめると以下のようになる。

(1)センサの配置情報を厳密に知る必要がなく、ある基準センサから他のセンサへの距離の上限のみを知ればよい。ため、不規則なセンサ配置を採用でき、さらに位置をキャリブレーションする必要が無い。(2)基底ベクトルから得られる情報をすべて用いてクラスタリングを行うためパーミュテーション問題をより正確に解決することができ、信号分離の性能が向上する。

- [0108] なお、本発明は上述の実施の形態に限定されるものではない。例えば、上述した各実施の形態では、一般化行列としてムーア・ペンローズ型一般化逆行列を用いたが他の一般化行列を利用する形態でもよい。

また、第1の実施の形態では、周波数正規化部142aの第1正規化部142aaが、基底ベクトル $A_p(f)$ の特定の1つの要素 $A_{qp}(f)$ を基準とし、式(15)によって当該基底ベクトル $A_p(f)$ の各要素 $A_{qp}(f)$ の偏角を正規化することとした。しかし、例えば、以下のような式に従って、第1正規化部142aaが、基底ベクトル $A_p(f)$ の特定の1つの要素 $A_{qp}(f)$ を基準として、当該基底ベクトル $A_p(f)$ の各要素 $A_{qp}(f)$ の偏角を正規化する構成であつてもよい。

- [0109] [数30]

$$A_{qp}'''(f) = |A_{qp}(f)| \exp\{j \cdot (\arg[A_{qp}(f) \cdot A_{qp}^*(f)])\} \quad \dots (27-1)$$

$$A_{qp}'''(f) = |A_{qp}(f)| \exp\{j \cdot (\arg[A_{qp}(f)] - \arg[A_{qp}(f)])\} \quad \dots (27-2)$$

$$A_{qp}'''(f) = |A_{qp}(f)| \exp\{j \cdot \phi(\arg[A_{qp}(f)/A_{qp}(f)])\} \quad \dots (27-3)$$

ただし、 \cdot^* は、 \cdot の複素共役である。また、 $\phi\{\cdot\}$ は関数であり、クラスタリング精度の観点から好ましくは単調増加関数であることが望ましい。

また、周波数正規化部142aが式(14)の代わりに、

[0110] [数31]

$$A_{qp}'(f) = \rho \cdot \frac{\arg[A_{qp}(f)/A_{qp}(f)]}{4fc^{-1}d} \quad \dots (28-1)$$

$$A_{qp}'(f) = \rho \cdot \frac{\arg[A_{qp}(f) \cdot A_{qp}^*(f)]}{4fc^{-1}d} \quad \dots (28-2)$$

$$A_{qp}'(f) = \rho \cdot \frac{\arg[A_{qp}(f)] - \arg[A_{qp}(f)]}{4fc^{-1}d} \quad \dots (28-3)$$

$$A_{qp}'(f) = \rho \cdot \frac{\phi(\arg[A_{qp}(f)/A_{qp}(f)])}{4fc^{-1}d} \quad \dots (28-4)$$

等の演算によって周波数正規化を行うこととしてもよい。ただし、 ρ は定数である(例えば、 $\rho=1$)。

また、上述の実施の形態では、ノルム正規化部142bにおいてノルムが1になるような正規化を行ったが、ノルムが1以外の規定値となるように正規化を行ってもよい。さらに、ノルム正規化部142bを設けず、ノルム正規化を行わない構成であってもよい。この場合、クラスタリング部143は、周波数正規化ベクトル $A_p'(f)$ のクラスタリングを行うことになる。しかし、周波数正規化ベクトル $A_p'(f)$ は、ノルムが統一されていない。そこで、この場合のクラスタリング基準は、ベクトルがノルムを含めて似ているかどうかではなく、ベクトルの方向のみが似ているかどうかになる。これは、類似度を用いた評価になる。類似度の1つとしてコサイン距離

$$\cos \theta = |A_p'^H(f) \cdot \eta_k| / (\|A_p'(f)\| \cdot \|\eta_k\|)$$

を例示できる。ここで θ は、周波数正規化ベクトル $A_p'(f)$ と、セントロイド η_k のベクトル

とがなす角度である。コサイン距離を用いる場合、クラスタリング部143は、コサイン距離の総和

[0111] [数32]

$$U_i = \sum_{A_p'(f) \in C_i} |A_p'^H(f) \cdot \eta_i| / (\|A_p'(f)\| \cdot \|\eta_i\|)$$

を最小値化するクラスタを生成する。なお、セントロイド η_k は、各クラスタのメンバの平均として算出する。

さらに、上記の第2の実施の形態では、周波数毎に順列の信頼度を評価し、順列の信頼度が低いと評価された周波数に対し、分離信号のエンベロープを用いて新たな順列を算出することとした。しかし、分離信号のエンベロープとクラスタの中心ベクトルと正規化基底ベクトルとを一度に統合して順列を算出することとしてもよい。

[0112] また、まず分離信号のエンベロープを用いて順列を算出し、その順列の信頼度を周波数毎に評価し、信頼度が低いと評価された周波数に対して、第1の実施の形態の手法を適用して新たな順列を算出することとしてもよい。

さらに、上記の第2の実施の形態では、第2の分離行列 $W'(f)$ を用いて分離信号のエンベロープを算出したが、第1の分離行列 $W(f)$ 或いはその行を並び替えた分離行列を用いてエンベロープを算出する構成としてもよい。

また、パラメータ d は、全てのセンサ q に対して同一であってもよいし、各センサ q に対応して複数の値を設定しておいてもよい。例えば、基準センサとセンサ q との距離を、それぞれのセンサ q に対応するパラメータ d の値としてもよい。

[0113] [第3の実施の形態(「第2の本発明」の例)]

次に、本発明における第3の実施の形態について説明する。

本形態は、前述の原理を用い、目的信号の方向に関する情報を知ることなく、複数の信号源から発せられた信号が混合された混合信号から目的信号を抽出する形態である。

<構成>

本形態の信号分離装置も、第1の実施の形態と同様、公知のノイマン型のコンピュータに信号分離プログラムが読み込まれることにより構成される。図16は、第3の実

施の形態における信号分離装置1001の構成を例示したブロック図である。

[0114] この図に例示するように、信号分離装置1001は、記憶領域1101～1114を有するメモリ1100、周波数領域変換部1120、信号分離部1130、目的信号選択部1140、時間周波数マスキング部1150(「分離信号生成部」に相当)、時間領域変換部1160、制御部1170及び一時メモリ1180を有する。ここでメモリ1100及び一時メモリ1180は、例えば、レジスタ10ac、補助記憶装置10fのデータ領域10fb或いはRAM10dのデータ領域10db(図2)の少なくとも一部に相当するが、特にこれに限定されない。また、周波数領域変換部1120、信号分離部1130、目的信号選択部1140、時間周波数マスキング部1150、時間領域変換部1160、制御部1170及び一時メモリ1180は、例えば、CPU10a(図2)にOSプログラムや信号分離プログラムが読み込まれることにより構成されるものである。

[0115] 図17Aは、図16における目的信号選択部1140の詳細構成を例示したブロック図であり、図17Bは、図17Aにおける基底ベクトルクラスタリング部1142の詳細構成を例示したブロック図である。

これらの図に例示するように、目的信号選択部1140は、逆行列算出部1141(「複素ベクトル生成部」に対応)、基底ベクトルクラスタリング部1142及び選択部1143を有している。また、基底ベクトルクラスタリング部1142は、周波数正規化部1142a(「正規化部」を構成)、ノルム正規化部1142b(「正規化部」を構成)、クラスタリング部1142c及び分散判定部1142dを有している。また、周波数正規化部1142aは、第1正規化部1142aa及び第2正規化部1142abを有している。

[0116] 図18Aは、図16における時間周波数マスキング部1150の詳細構成を例示したブロック図であり、図18Bは、図18Aにおけるマスク生成部1151の詳細を例示したブロック図である。

これらの図に例示するように、時間周波数マスキング部1150は、マスク生成部1151及びマスキング部1152を有している。また、マスク生成部1151は、白色化行列生成部1151a、白色化部1151b、角度計算部1151c及び関数演算部1151dを有している。

[0117] なお、図16～図18における実線の矢印は実際のデータの流れを示し、破線の矢

印は理論上の情報の流れを示すが、制御部1170及び一時メモリ1180に出入りするデータの流れは省略してある。また、信号分離装置1001は制御部1170の制御のもと各処理を実行する。また、特に示さない限り、制御部1170は、必要なデータを一時メモリ1180に読み書きしながら各処理を実行していく。

<処理>

次に、本形態における信号分離装置1001の処理について説明する。

- [0118] 空間内にN個の信号源 k ($k \in \{1, 2, \dots, N\}$)が存在し、それらの信号 $s_k(t)$ (t はサンプリング時刻)が混合され、M個のセンサ q ($q \in \{1, 2, \dots, M\}$)で混合信号 x_q として観測されるとする。本形態では、混合信号 $x_1(t), \dots, x_M(t)$ のみから、何れかの信号源から発せられた目的信号を抽出してその他の妨害信号を抑圧した信号 $y(t)$ を算出する。なお、信号源の数 N はセンサの数 M よりも多くても少なくてもよく、また同数であってもよい。さらに、信号源の数 N を事前に知る必要もなく、また信号源の数が数えられないような状況であっても構わない。

[0119] [処理の全体]

図19は、第3の実施の形態における信号分離処理の全体を説明するためのフローチャートである。以下、この図を用いて本形態における信号分離処理の全体を説明する。

まず、前処理において、M個のセンサで観測された時間領域の混合信号 $x_q(t)$ ($q \in \{1, \dots, M\}$)をメモリ1100の記憶領域1101に格納しておく。信号分離処理が開始されると、まず、周波数領域変換部1120において、メモリ1100の記憶領域1101から時間領域の混合信号 $x_q(t)$ を読み出す。そして、周波数領域変換部1120は、これらを短時間フーリエ変換などにより周波数領域の混合信号 $X_q(f, \tau)$ に変換し、メモリ1100の記憶領域1102に格納する(ステップS101)。

- [0120] 次に、信号分離部1130において、メモリ1100の記憶領域1102から周波数領域の混合信号 $X_q(f, \tau)$ を読み出す。この例の信号分離部1130は、読み出した混合信号 $X_q(f, \tau)$ を要素とする混合信号ベクトル $X(f, \tau) = [X_1(f, \tau), \dots, X_M(f, \tau)]^T$ に、独立成分分析(ICA)を適用し、周波数 f 毎にM行M列の分離行列 $W(f) = [W_1(f), \dots, W_M(f)]^H$ ($*$ ^Hは行列 $*$ の複素共役転置行列)と分離信号ベクトル

$$Y(f, \tau) = W(f) \cdot X(f, \tau) \quad \dots (30)$$

とを算出する(ステップS102)。なお、算出された分離行列 $W(f)$ はメモリ1100の記憶領域1103に格納される。また、分離信号ベクトル $Y(f, \tau) = [Y_1(f, \tau), \dots, Y_M(f, \tau)]^T$ を構成する分離信号 $Y_p(f, \tau)$ ($p \in \{1, \dots, M\}$)は、記憶領域1107に格納される。ステップS102の詳細については後述する。

[0121] 次に、目的信号選択部1140において、メモリ1100の記憶領域1103から分離行列 $W(f)$ を読み出し、その一般化逆行列の列である基底ベクトルを正規化し、その正規化された基底ベクトルをクラスタリングする。そして、目的信号選択部1140は、そのクラスタの分散を指標として、メモリ1100の記憶領域1107の分離信号から目的信号を含む選択信号 $Y_{10}(f, \tau)$ とそれに対応する基底ベクトル $A_{10}(f)$ とを周波数 f 毎に選択し、メモリ1100の記憶領域1111に格納する(ステップS103)。なお、本形態では、センサに近く、そのためセンサで観測されるパワーがその他の源信号と比べて支配的であり、さらに情報として有用である信号を目的信号として抽出する。このステップS103の詳細については後述する。

[0122] 次に、時間周波数マスキング部1150において、メモリ1100の記憶領域1102から周波数領域の混合信号 $X_q(f, \tau)$ を読み出し、記憶領域1104から選択信号 $Y_{10}(f, \tau)$ に対応する基底ベクトル $A_{10}(f)$ を読み出し、これらを用いて時間周波数マスク $M(f, \tau)$ を生成して記憶領域1112に格納する(ステップS104)。なお、ステップS104の処理(時間周波数マスキング部1150の処理)の詳細については後述する。

次に、時間周波数マスキング部1150において、メモリ1100の記憶領域1107から目的信号選択部1140で選択された選択信号 $Y_{10}(f, \tau)$ を読み出し、記憶領域1112から時間周波数マスク $M(f, \tau)$ を読み出す。そして、時間周波数マスキング部1150は、この時間周波数マスク $M(f, \tau)$ を選択信号 $Y_{10}(f, \tau)$ に適用し、選択信号 $Y_{10}(f, \tau)$ に残存する妨害信号成分をより一層抑圧したマスク処理選択信号 $Y_{10}'(f, \tau)$ を生成し、メモリ1100の記憶領域1113に格納する(ステップS105)。なお、このステップS105の処理の詳細(時間周波数マスキング部1150の処理)については後述する。

[0123] 最後に、時間領域変換部1160において、メモリ1100の記憶領域1113から、周波数領域における選択された分離信号 $Y_{10}'(f, \tau)$ を読み出し、それに短時間逆フーリ

エ変換などを施して時間領域の分離信号 $y(t)$ を生成し、メモリ1100の記憶領域1114に格納する(ステップS106)。

[ステップS102の処理(信号分離部1130の処理)の詳細]

前述のように、この例の信号分離部1130は、独立成分分析(ICA)によって、混合信号ベクトル $X(f, \tau)=[X_1(f, \tau), \dots, X_M(f, \tau)]^T$ からM行M列の分離行列 $W(f)=[W_1(f), \dots, W_M(f)]^H$ と分離信号ベクトル $Y(f, \tau)=[Y_1(f, \tau), \dots, Y_M(f, \tau)]^T$ とを算出する(ステップS102)。独立成分分析(ICA)は、分離信号ベクトル $Y(f, \tau)=[Y_1(f, \tau), \dots, Y_M(f, \tau)]^T$ の各要素が互いに独立になるように分離行列 $W(f)$ を算出する手法であり、そのアルゴリズムとしては、非特許文献4に記載されているものなど様々なものが紹介されている。なお、独立成分分析(ICA)は、パワーが小さく、ガウス性が高い妨害信号よりも、パワーが大きく、非ガウス性が高い本形態の目的信号をより有利に分離抽出できる手法である。

[0124] [ステップS103の処理(目的信号選択部1140の処理)の詳細]

独立成分分析(ICA)では、信号の独立性に着目してそれらの分離を行うため、得られる分離信号 $Y_p(f, \tau)$ には順序の任意性がある。その順序が入れ替わっても独立性が保たれるからである。従って、目的信号に対応する分離信号を周波数毎に選択しなければならない。目的信号選択部1140では、以下に示す処理により、この選択を行う。

図20は、第3の実施の形態における目的信号選択部1140の処理の詳細を説明するためのフローチャートである。以下では、この図を用い、目的信号選択部1140の処理の詳細を説明する。

[0125] まず、逆行列算出部1141において、周波数毎にメモリ1100の記憶領域1103からM行M列の分離行列 $W(f)$ を読み出し、それらの逆行列

$$W(f)^{-1}=[A_1(f), \dots, A_M(f)] \text{ (ただし、その列 } A_p(f)=[A_{1p}(f), \dots, A_{Mp}(f)]^T \text{)} \cdots (31)$$

を算出する。

ここで、式(30)の両辺に式(31)を掛け合わせることで、周波数領域の混合信号 $X(f, \tau)$ の分解

[0126] [数33]

$$X(f, \tau) = \sum_{p=1}^M A_p(f) Y_p(f, \tau) \quad \dots (32)$$

を得ることができる。この $A_p(f)$ は基底ベクトルであり、それぞれが各周波数の分離信号 $Y_p(f, \tau)$ 毎に対応する。式(31)のように算出された基底ベクトル $A_p(f)$ はメモリ1100の記憶領域1104に格納される(ステップS111)。

次に、基底ベクトルクラスタリング部1142において、全ての基底ベクトル $A_p(f)$ ($p=1, \dots, M, f=0, f/L, \dots, f(L-1)/L$)を正規化する。この正規化は、複数の信号源から発生された信号の畳み込み混合を所定のモデル(例えば、近距離場モデル)に近似した場合に、正規化された基底ベクトル $A_p(f)$ が信号源の位置のみに依存したクラスタを形成するように行われる。この例では、第1の実施の形態と同様な周波数正規化とノルム正規化とを行う。

[0127] 周波数正規化は、基底ベクトルクラスタリング部1142の周波数正規化部1142a(図17B)において行われる。すなわち、周波数正規化部1142aにおいて、メモリ1100の記憶領域1104の基底ベクトル $A_p(f)$ ($p=1, \dots, M, f=0, f/L, \dots, f(L-1)/L$)を読み出し、これらを周波数に依存しない周波数正規化ベクトル $A'_p(f)$ に正規化してメモリ1100の記憶領域1105に格納する(ステップS112)。なお、この正規化は基底ベクトル $A_p(f)$ の要素 $A_{fp}(f)$ 毎に行われる(詳細は後述)。ノルム正規化は、基底ベクトルクラスタリング部1142のノルム正規化部1142b(図17B)において行われる。すなわち、ノルム正規化部1142bにおいて、メモリ1100の記憶領域1105から周波数正規化ベクトル $A'_p(f)$ を読み出し、これらをノルムが規定値(この例では1)をとる正規化基底ベクトル $A''_p(f)$ に正規化してメモリ1100の記憶領域1106に格納する(ステップS113)。なお、この正規化は周波数正規化ベクトル $A'_p(f)$ 毎に行われる(詳細は後述)。

[0128] 基底ベクトルの正規化が終了すると、次にクラスタリング部1142c(図17B)において、正規化基底ベクトル $A''_p(f)$ によって形成されるM個のクラスタ C_i ($i \in \{1, \dots, M\}$)を同定する。この例の場合、クラスタリング部1142cにおいて、メモリ1100の記憶領域1106から正規化基底ベクトル $A''_p(f)$ を読み出し、これらをクラスタリングしてM個のクラスタ C_i ($i=1, \dots, M$)を生成し、各クラスタ C_i を特定する情報(例えば、各クラスタに属する正規化基底ベクトル $A''_p(f)$ を示す情報)と各クラスタ C_i のセントロイド(中心ベクトル)

η_i とをそれぞれメモリ1100の記憶領域1109, 1110に格納する(ステップS114)。なお、このクラスタリングは、例えば、各クラスタ C_i の要素(正規化基底ベクトル $A_v''(f)$)と各クラスタ C_i のセントロイド η_i との間の二乗和 U_i の総和 U

[0129] [数34]

$$U = \sum_{i=1}^M U_i \quad \dots (33)$$

$$U_i = \sum_{A_v''(f) \in C_i} \|A_v''(f) - \eta_i\|^2$$

を最小値化することを基準に行われる。この最小値化は、例えば、非特許文献6などで解説されているk-meansクラスタリングを用いることによって効果的に行うことができる。なお、クラスタ C_i のセントロイド η_i は、

[0130] [数35]

$$\eta_i = \frac{\sum_{A_v''(f) \in C_i} A_v''(f) / |C_i|}{\left\| \sum_{A_v''(f) \in C_i} A_v''(f) / |C_i| \right\|} \quad \dots (34)$$

によって計算される。ここで $|C_i|$ はクラスタ C_i における要素(正規化基底ベクトル $A_v''(f)$)の数であり、 $\|*\|$ はベクトル $*$ のノルムである。また、ここでは、距離としてユークリッド距離の自乗を用いているが、これを一般化したミンコフスキー距離などを用いてもよい。

M個のクラスタ C_i が得られると、次に分散判定部1142d(図17B)において、目的信号に対応するクラスタを選択し、選択したクラスタを示す選択情報 $I(f)$ をメモリ1100の記憶領域1111に格納する(ステップS115)。なお本形態では、クラスタの分散 $U_i/|C_i|$ を指標に、目的信号を含む分離信号が選択される。すなわち、本形態における基底ベクトルの正規化は、複数の信号源から発せられた信号の畳み込み混合を所定のモデルに近似した場合に、正規化された基底ベクトルが信号源の位置のみに依存したクラスタを形成するように行われる。しかし、実環境では、このようなモデルに反映されない様々な因子が存在し、このような実環境とモデルとの乖離は、信号源からセンサまでの距離が大きいほど大きくなる。例えば、近距離場モデルでは、信号が壁などに反射することによる反射波成分が考慮されておらず、直接波成分に対する反射

波成分の割合は、信号源がセンサから離れるに従って大きくなる。そのため、信号がセンサから離れるに従ってモデルが成り立ちにくくなる。その結果、センサに近い信号ほど実環境に近い条件で正規化ができ、実環境とモデルとの乖離に起因するクラスタの分散を小さくできる。本形態ではセンサに近い信号を目的信号とするため、分散の小さいクラスタを目的信号に対応するクラスタとして選択すればよいことになる。なお、この選択手順(ステップS115)の詳細については後述する。

[0131] 各周波数 f に対する選択情報 $I(f)$ が算出されると、次に、各周波数 f における選択信号 $Y_{10}(f, \tau)$ とそれに対応する基底ベクトル $A_{10}(f)$ とを選択する。すなわち、まず、選択部1143において、メモリ1100の記憶領域1111から選択情報 $I(f)$ を読み出す。そして、選択部1143は、この選択情報 $I(f)$ に対応する分離信号を選択信号 $Y_{10}(f, \tau)$ として記憶領域1107から読み出し、それに対応する基底ベクトル $A_{10}(f)$ を記憶領域1104から読み出し、これらを記憶領域1111に格納する(ステップS116)。

[0132] 次に、ステップS112, 113(図20)の正規化の詳細を示す。

[ステップS112(周波数正規化)の詳細]

図21Aは、ステップS112の周波数正規化の詳細を説明するためのフローチャートである。

まず、制御部1170(図16)においてパラメータ p に1を代入し、これを一時メモリ1180に格納する(ステップS121)。また、制御部1170においてパラメータ q に1を代入し、これを一時メモリ1180に格納する(ステップS122)。次に、周波数正規化部1142a(図17B)において、メモリ1100の記憶領域1108から前述のパラメータ d, c, Q を読み出し、記憶領域1104から基底ベクトル $A_p(f)$ の要素 $A_{qp}(f)$ を読み出し、一時メモリ1180からパラメータ p, q を読み出す。そして、この周波数正規化部1142aは、基底ベクトル $A_p(f)$ の要素 $A_{qp}(f)$ に対し、

[0133] [数36]

$$A_{qp}'(f) = A_{qp}(f) \exp \left[j \frac{\arg[A_{qp}(f)/A_{Qp}(f)]}{4fc^{-1}d} \right] \quad \dots (35)$$

の演算を行い、当該演算結果 $A_{qp}'(f)$ を周波数正規化ベクトル $A_p'(f)$ の各要素 $A_{qp}'(f)$ としてメモリ1100の記憶領域1105に格納する(ステップS123)。なお、 $\arg[\cdot]$ は偏

角を意味し、expはネピア数を意味し、jは虚数単位を意味する。より詳細には、前述の式(15)(16)の演算によって、この正規化を行う。

次に、制御部1170において、一時メモリ1180に格納されたパラメータqが $q=M$ を満たすか否かを判断する(ステップS124)。ここで $q=M$ でなければ、制御部1170は、 $q+1$ の演算結果を新たなパラメータqの値とし、これを一時メモリ1180に格納し(ステップS125)、処理をステップS123へ戻す。一方、 $q=M$ であれば、制御部1170は、さらに $p=M$ を満たすか否かを判断する(ステップS126)。

[0134] ここで $p=M$ でなければ、制御部1170は、 $p+1$ の演算結果を新たなパラメータpの値とし、これを一時メモリ1180に格納し(ステップS127)、処理をステップS122へ戻す。一方、 $p=M$ であれば、制御部1170は、ステップS112の処理を終了させる(ステップS112(周波数正規化)の詳細の説明終わり)。

[ステップS113(ノルム正規化)の詳細]

図21Bは、ステップS113のノルム正規化の詳細を説明するためのフローチャートである。

[0135] まず、制御部1170において、パラメータpに1を代入し、これを一時メモリ1180に格納する(ステップS131)。次に、ノルム正規化部1142bにおいて、メモリ1100の記憶領域1105から周波数正規化ベクトル $A_p'(f)$ の各要素 $A_{pq}'(f)$ を読み出し、

[0136] [数37]

$$\|A_p'(f)\| = \sqrt{\sum_{q=1}^M (A_{pq}'(f))^2} \quad \dots (38)$$

の演算を行って、周波数正規化ベクトル $A_p'(f)$ のノルム $\|A_p'(f)\|$ を求め、周波数正規化ベクトル $A_p'(f)$ 及びそのノルム $\|A_p'(f)\|$ を一時メモリ1180に格納する(ステップS132)。

次に、ノルム正規化部1142bにおいて、一時メモリ1180から周波数正規化ベクトル $A_p'(f)$ 及びそのノルム $\|A_p'(f)\|$ を読み出し、

$$A_p''(f) = A_p'(f) / \|A_p'(f)\| \quad \dots (39)$$

の演算を行って正規化基底ベクトル $A_p''(f)$ を求め、これをメモリ1100の記憶領域1106に格納する(ステップS133)。その後、制御部1170において、一時メモリ1180

に格納されたパラメータ p が $p=M$ を満たすか否かを判断する(ステップS134)。ここで $p=M$ でなければ、制御部1170は、 $p+1$ の演算結果を新たなパラメータ p の値とし、これを一時メモリ1180に格納し(ステップS135)、ステップS132の処理に戻す。一方、 $p=M$ であれば、制御部1170は、ステップS113の処理を終了させる。なお、正規化基底ベクトル $A_p''(f)$ がクラスタを形成する理由は、第1の実施の形態で示した通りである(ステップS113(ノルム正規化)の詳細の説明終わり)。

- [0137] このように生成された正規化基底ベクトル $A_p''(f)$ は、第1の実施の形態で示した通り、周波数に依存せず、信号源の位置のみに依存するベクトルとなる。

[選択信号の選択手順(ステップS115)の詳細]

次に、前述した選択信号の選択手順(ステップS115)の詳細を例示する。

クラスタの選択手順1:

第1の例は、最も分散が小さいクラスタを目的信号に対応するクラスタとして選択するものである。図22は、この第1の例を説明するためのフローチャートである。

- [0138] この場合、まず分散判定部1142d(図17B)において、メモリ1100の記憶領域1109からクラスタ $C_i (i \in \{1, \dots, M\})$ を特定する情報を読み出し、さらに記憶領域1106, 1110から、それぞれ正規化基底ベクトル $A_p''(f) \in C_i$ 及びセントロイド η_i を読み出す。そして、分散判定部1142dは、各 i について、式(33)のように U_i を求め、 C_i に属する要素(正規化基底ベクトル $A_p''(f)$)の数をカウントすることにより $|C_i|$ を求め、クラスタ C_i の分散 $U_i/|C_i|$ を算出して一時メモリ1180に格納する。そして、分散判定部1142dは、この一時メモリ1180に格納された分散 $U_i/|C_i|$ から最小なものを選択し、それに対応するクラスタを示す情報をクラスタ選択情報

$$i = \operatorname{argmin}_i U_i/|C_i| \dots (40)$$

として一時メモリ1180に格納する(ステップS141)。なお、式(40)における $\operatorname{argmin}_i *$ は、 $*$ を最小にする i を意味する。

- [0139] 次に、制御部1170(図16)において、パラメータ f に0を代入し、これを一時メモリ1180に格納する(ステップS142)。

次に、分散判定部1142dは、一時メモリ1180からクラスタ選択情報 i を読み出し、メモリ1100の記憶領域1110からクラスタ選択情報 i に対応するセントロイド η_i を読

み出す。また、分散判定部1142dは、メモリ1100の記憶領域1106から正規化基底ベクトル $A_p''(f)$ ($p \in \{1, \dots, M\}$)を読み出す。そして、分散判定部1142dは、周波数 f 毎に選択情報

$$I(f) = \underset{p}{\operatorname{argmin}} \| A_p''(f) - \eta \| ^2 \dots (41)$$

を算出し、記憶領域1111に格納する(ステップS143)。

- [0140] 次に、制御部1170において、一時メモリ1180からパラメータ f を読み出し、 $f = (L-1) \cdot f_s/L$ であるか否かを判断する(ステップS144)。ここで $f = (L-1) \cdot f_s/L$ でなければ、制御部1170は、このパラメータ f に f_s/L を加算した値を新たな f として一時メモリ1180に格納し(ステップS145)、処理をステップS143に戻す。一方、 $f = (L-1) \cdot f_s/L$ であれば、制御部1170は、ステップS115の処理を終了させる。

クラスタの選択手順2:

第2の例は、予め定められたしきい値よりも分散が小さいクラスタを目的信号に対応するクラスタとして選択するものである。なお、このしきい値は、例えば実験結果又は経験に基づき定められる値であり、事前にメモリ1100に格納しておくものである。

- [0141] この場合、ステップS141(図22)の処理の代わりに、分散判定部1142dにおいて、まず各クラスタの分散 $U_i/|C_i|$ を公知の並び替えアルゴリズムによって昇順或いは降順に並び替える。そして、分散判定部1142dは、メモリ1100に格納されたしきい値を読み取り、分散 $U_i/|C_i|$ がしきい値よりも小さくなるクラスタを選択し、選択したクラスタに対応する添字 i の集合をクラスタ選択情報 ι として一時メモリ1180に格納する。それ以外は、クラスタの選択手順1と同様である。

クラスタの選択手順3:

第3の例は、最も分散の小さいクラスタだけではなく、最も分散が小さいクラスタから所定番目に分散が小さいクラスタまで(例えば、最も分散の小さいクラスタから3番目に分散が小さいクラスタまで)を目的信号に対応するクラスタとして選択するものである。

- [0142] この場合、ステップS141(図22)の処理の代わりに、分散判定部1142dにおいて、まず各クラスタの分散 $U_i/|C_i|$ を、公知の並び替えアルゴリズムによって昇順或いは降順に並び替える。そして、分散判定部1142dは、最も分散 $U_i/|C_i|$ が小さいクラスタか

ら所定番目に分散 $U_i/|C_i|$ が小さいクラスタまでを選択する。この選択後、分散判定部1142dは、選択したクラスタに対応する添字 i の集合をクラスタ選択情報 i として一時メモリ1180に格納する。それ以外は、クラスタの選択手順1と同様である。

なお、クラスタの選択手順1の代わりに分散の小ささが2番目以降のいずれかのクラスタを選択することとしてもよく、上述したクラスタの選択手順の一部を組み合わせた手順を適用してもよい(ステップS115の説明終わり/ステップS103の処理(目的信号選択部1140の処理)の詳細の説明終わり)。

[0143] [時間周波数マスキング部1150の処理(ステップS104, S105)の詳細]

次に、時間周波数マスキング部1150の処理について説明する。前述のように、時間周波数マスキング部1150は、目的信号選択部1140で選択された選択信号 $Y_{i(f)}(f, \tau)$ に残存する妨害信号成分をさらに抑圧するものである。以下ではまず、なぜ選択信号 $Y_{i(f)}(f, \tau)$ に妨害信号が残存するのかについて説明する。

前述した式(30)を選択信号のみに関して書くと、

$$Y_{i(f)}(f, \tau) = W_{i(f)}^H(f) \cdot X(f, \tau) \quad \cdots (42)$$

となる。また、これに式(4)を代入して周波数 f を省略して書くと、

[0144] [数38]

$$Y_i(\tau) = W_i^H \cdot H_i \cdot S_i(\tau) + \sum_{k=1, \dots, I-1, I+1, \dots, N} W_i^H \cdot H_k \cdot S_k(\tau) \quad \cdots (43)$$

と表記できる。もし $N \leq M$ であれば、独立成分分析(ICA)により、 $W_i^H \cdot H_k = 0, \forall k \in \{1, \dots, I-1, I+1, \dots, N\}$ を満たすような W_i を設定することが可能である。そのため、式(43)の第二項は0になる。しかし、より一般的に信号源数 N がセンサ数 M よりも多い場合には、 $W_i^H \cdot H_k \neq 0, \forall k \in \kappa$ となる $\kappa \subseteq \{1, \dots, I-1, I+1, \dots, N\}$ が存在してしまう。その場合、選択信号 $Y_i(f)$ は不必要な残留成分(妨害信号の残留成分)

[0145] [数39]

$$\sum_{k \in \kappa} W_i^H \cdot H_k \cdot S_k(\tau)$$

を含んでしまう(以後 f を省略せず)。

時間周波数マスキング部1150の目的は、この選択信号 $Y_i(f, \tau)$ に含まれる不必要

な残留成分を抑圧し、選択信号 $Y_1(f, \tau)$ よりも妨害信号の残留成分が少ないマスク処理選択信号 $Y'_1(f, \tau)$ を生成することである。そのために、時間周波数マスキング部1150は、まずマスク生成部1151(図18)において、妨害信号の残留成分を有する時間周波数スロットに対して小さな値を採り、妨害信号の残留成分が少ない時間周波数スロットに対して大きな値を採る時間周波数マスク $0 \leq M(f, \tau) \leq 1$ を生成する。そして、マスキング部1152において、

$$Y'_{10}(f, \tau) = M(f, \tau) \cdot Y_{10}(f, \tau) \quad \cdots (44)$$

に従ってマスキング処理を行い、マスク処理選択信号 $Y'_{10}(f, \tau)$ を出力する。以下、この詳細について説明する。

[0146] [ステップS104(マスク生成部1151)の処理の詳細]

図23は、図19におけるステップS104の詳細を説明するためのフローチャートである。以下、この図を用い、ステップS104(マスク生成部1151)の処理の詳細を説明する。

この例のマスク生成部1151は、周波数領域の混合信号ベクトル $X(f, \tau)$ を白色化する空間(白色化空間)において、混合信号ベクトル $X(f, \tau)$ と選択信号に対応する基底ベクトル $A_{10}(f)$ とのなす角度 $\theta_{10}(f, \tau)$ を求め、その角度 $\theta_{10}(f, \tau)$ に基づき、時間周波数マスクを生成する。なお、白色化とは、混合信号ベクトル $X(f, \tau)$ を線形に変形し、その共分散行列が単位行列と等しくなるようにすることである。

[0147] そのためにまず、白色化行列生成部1151aにおいて、周波数領域の混合信号 $X(f, \tau)$ を用い、混合信号ベクトル $X(f, \tau)$ を白色化空間に移す白色化行列 $V(f)$ を生成する(ステップS151)。この例の場合、白色化行列生成部1151aは、メモリ1100の記憶領域1102から混合信号 $X(f, \tau)$ を読み出し、 $R(f) = \langle X(f, \tau) \cdot X(f, \tau)^H \rangle$ とした場合における $V(f) = R(f)^{-1/2}$ を白色化行列 $V(f)$ として算出し、これを記憶領域1112に格納する。なお、 $\langle * \rangle$ はベクトル $*$ を時間平均したベクトルを意味し、 $*^H$ は $*$ の複素共役転置を意味し、 $R^{-1/2}$ は、 $R^{-1/2} \cdot R \cdot (R^{-1/2})^H = I$ (I は単位行列)を満たす行列を意味する。また、白色化行列 $V(f)$ を算出する代表的な方法として、 $R(f)$ を固有値分解 $R(f) = E(f) \cdot D(f) \cdot E(f)^H$ ($E(f)$ はユニタリ行列、 $D(f)$ は対角行列)し、 $V(f) = D(f)^{-1/2} \cdot E(f)^H$ の演算を行う方法を例示できる。なお、 $D(f)^{-1/2}$ は、対角行列 $D(f)$ の各要素の(-

1/2)乗した対角行列と等価となるため、対角行列D(f)の各要素の(-1/2)乗することによって算出できる。

[0148] 次に、白色化部1151bにおいて、この白色化行列V(f)を用い、混合信号ベクトルX(f, τ)を白色化空間に写像した白色化混合信号ベクトルZ(f, τ)と、基底ベクトルA₁₀(f)を白色化空間に写像した白色化基底ベクトルB₁₀(f)とを算出する(ステップS152)。この例の場合、まず白色化部1151bにおいて、メモリ1100の記憶領域1102から混合信号X_q(f, τ)を、記憶領域1111から選択信号Y₁₀(f, τ)に対応する基底ベクトルA₁₀(f)を、記憶領域1112から白色化行列V(f)をそれぞれ読み出す。そして、白色化部1151bは、Z(f, τ)=V(f)・X(f, τ)の演算により白色化混合信号ベクトルZ(f, τ)を算出し、B₁₀(f)=V(f)・A₁₀(f)の演算により白色化基底ベクトルB₁₀(f)を算出し、これらをメモリ1100の記憶領域1112に格納する。

[0149] 次に、角度計算部1151cにおいて、白色化混合信号ベクトルZ(f, τ)と白色化基底ベクトルB₁₀(f)とがなす角度 θ_{10} (f, τ)を時間周波数毎に算出する(ステップS153)。この例の場合、まず角度計算部1151cにおいて、メモリ1100の記憶領域1112から白色化混合信号ベクトルZ(f, τ)と白色化基底ベクトルB₁₀(f)とを読み出す。そして、角度計算部1151cは、

$$\theta_{10}(f, \tau) = \cos^{-1}(|B_{10}^H(f) \cdot Z(f, \tau)| / \|B_{10}(f)\| \cdot \|Z(f, \tau)\|) \quad \dots (45)$$

の演算により、各時間周波数スロットに対して角度 θ_{10} (f, τ)を算出し、記憶領域1112に格納する。なお、式(45)における|*|はベクトル*の絶対値を意味し、||*||はベクトル*のノルムを意味する。

[0150] 次に、関数演算部1151dにおいて、角度 θ_{10} (f, τ)を要素とする関数である時間周波数マスクM(f, τ)を生成する(ステップS154)。この例の場合、関数演算部1151dにおいて、まず、メモリ1100の記憶領域1108から実数パラメータ θ_T, g を、記憶領域1112から角度 θ_{10} (f, τ)を、それぞれ読み出す。そして、関数演算部1151dは、ロジスティック関数

$$M(\theta(f, \tau)) = 1 / (1 + e^{g \cdot (\theta(f, \tau) - \theta_T)}) \quad \dots (46)$$

を時間周波数マスクM(f, τ)として算出する。なお、実数パラメータ θ_T, g は、それぞれ時間周波数マスクM(f, τ)の転移点と傾き度を指定するパラメータであり、前処理

によって記憶領域1108に格納しておいたものである。図24Aに、二種類の実数パラメータ θ_T , g に対し、式(46)に従って算出された時間周波数マスク $M(f, \tau)$ を例示する。この図に例示するように、実数パラメータ θ_T が小さいほど、時間周波数マスク $M(f, \tau)$ が大きな値(この例では1)をとる領域が狭くなる。これは実数パラメータ θ_T が小さいほど、マスク処理選択信号 $Y_{100}'(f, \tau)$ に残存する妨害信号成分が少なくなるが、その反面、マスク処理選択信号 $Y_{100}'(f, \tau)$ が不自然なものとなるといった傾向に現れる。例えば、音声信号を目的信号とする場合、実数パラメータ θ_T を小さくするほどミュージカルノイズ(musical noise)が大きくなる。また、実数パラメータ g が大きくなるほど時間周波数マスク $M(f, \tau)$ の波形(大きな値(この例では1)から小さな値(この例では0)への遷移)は急峻になる。ここで、できるだけマスク処理選択信号 $Y_{100}'(f, \tau)$ に残存する妨害信号成分を減らし、マスク処理選択信号 $Y_{100}'(f, \tau)$ を自然な信号とするためには実数パラメータ g を小さくし、時間周波数マスク $M(f, \tau)$ の波形を滑らかなものとするのが望ましい。

[0151] なお、これらの実数パラメータ θ_T , g を周波数毎に設定しておく構成としてもよい。

また、さらに実数パラメータ α を導入し、ロジスティック関数

$$M(\theta(f, \tau)) = \alpha / (1 + e^{g \cdot (\theta(f, \tau) - \theta_T)}) \quad \dots (47)$$

を時間周波数マスク $M(f, \tau)$ としてもよい。さらに、角度 $\theta_{100}(f, \tau)$ が0に近い領域において大きな値を採り、角度 $\theta_{100}(f, \tau)$ が大きな領域において小さな値を採る $0 \leq M(\theta(f, \tau)) \leq 1$ の関数であれば、他の関数を時間周波数マスク $M(f, \tau)$ として用いてもよい(ステップS104(マスク生成部1151)の処理の詳細の説明終わり)。

[0152] [ステップS105(マスキング部1152)の処理の詳細]

マスキング部1152は、メモリ1100の記憶領域1111から選択信号 $Y_{100}(f, \tau)$ を、記憶領域1112から時間周波数マスク $M(f, \tau)$ をそれぞれ読み出す。そして、マスキング部1152は、

$$Y_{100}'(f, \tau) = M(f, \tau) \cdot Y_{100}(f, \tau) \quad \dots (48)$$

の演算によってマスク処理選択信号 $Y_{100}'(f, \tau)$ を求め、これをメモリ1100の記憶領域1113に格納する(ステップS105(マスキング部1152)の処理の詳細の説明終わり)。

[0153] [時間周波数マスクングの効果]

次に、上述した時間周波数マスク $M(f, \tau)$ によるマスクングの効果について説明する。

信号源のスパース性が高く、信号源 $S_k(f, \tau)$ がゼロに近くなる可能性が高いなら、式(4)は、

[0154] [数40]

$$X(f, \tau) \approx H_k(f) \cdot S_k(f, \tau), k \in \{1, \dots, N\} \quad \dots (49)$$

として近似される。なお、式(49)における k は各信号源に対応する添字であり、時間周波数位置 (f, τ) 毎に決まる。よって、目的信号が唯一アクティブ或いはそれに近い時間周波数位置 (f, τ) では、白色化混合信号ベクトル $Z(f, \tau)$ は、

[0155] [数41]

$$Z(f, \tau) \approx V(f) \cdot H_{I(f)}(f) \cdot S_{I(f)}(f, \tau) \approx V(f) \cdot A_{I(f)}(f) \cdot Y_{I(f)}(f, \tau)$$

と近似できる。ここで、 $Y_{I(f)}(f, \tau)$ はスカラー値である。また、前述のように白色化基底ベクトル $B_{I(f)}(f)$ は、

$$B_{I(f)}(f) = V(f) \cdot A_{I(f)}(f) \quad \dots (50)$$

である。以上より、目的信号が唯一アクティブ或いはそれに近い時間周波数位置 (f, τ) では、白色化混合信号ベクトル $Z(f, \tau)$ と白色化基底ベクトル $B_{I(f)}(f)$ とがなす角度 $\theta_{I(f)}(f, \tau)$ は0に近くなることが分かる。そして、前述のように時間周波数マスク $M(f, \tau)$ は、角度 $\theta_{I(f)}(f, \tau)$ が0に近い領域において大きな値を採るのである。よって、時間周波数マスク $M(f, \tau)$ は、目的信号が唯一アクティブ或いはそれに近い時間周波数位置 (f, τ) における選択信号 $Y_{I(f)}(f, \tau)$ をマスク処理選択信号 $Y_{I(f)}'(f, \tau)$ として抽出するものといえる(式(48)参照)。

[0156] 一方、仮に $I(f)=1$ とおいた場合、目的信号が殆どアクティブでない時間周波数位置 (f, τ) では、白色化混合信号ベクトル $Z(f, \tau)$ は、

[0157] [数42]

$$Z(f, \tau) \approx \sum_{k=2}^N V(f) \cdot H_k(f) \cdot S_k(f, \tau) \quad \dots (51)$$

と近似できる。ここで、信号源の N が、センサの数 M と同等かそれより少ないならば、白色化空間でのベクトル $V(f) \cdot H_1(f), \dots, V(f) \cdot H_k(f)$ は互いに直交する。なお、式(51)の $S_k(f, \tau)$ はスカラー値である。これより、白色化混合信号ベクトル $Z(f, \tau)$ と白色化基底ベクトル $B_{10}(f)$ とがなす角度 $\theta_{10}(f, \tau)$ が大きくなることがわかる。また $N > M$ でも、白色化基底ベクトル $B_1(f) (I(f)=1)$ は、目的信号以外のベクトル $V(f) \cdot H_2(f), \dots, V(f) \cdot H_k(f)$ とは大きい角度を持つ傾向がある。以上より、目的信号が殆どアクティブでない時間周波数位置 (f, τ) では、角度 $\theta_{10}(f, \tau)$ が大きな値を採ることが分かる。そして、時間周波数マスク $M(f, \tau)$ は、角度 $\theta_{10}(f, \tau)$ が0から離れた領域において小さな値を採るのであるから、目的信号が殆どアクティブでない時間周波数位置 (f, τ) における選択信号 $Y_{10}(f, \tau)$ をマスク処理選択信号 $Y_{10}'(f, \tau)$ から排除するものといえる(式(28)参照)。

[0158] 以上より、時間周波数マスク $M(f, \tau)$ による時間周波数マスキング処理は、選択信号 $Y_{10}(f, \tau)$ に残存する妨害信号成分をより一層抑圧するものであることがわかる。

なお、この時間周波数マスキング処理は、音声や音楽などスパース性を持つ信号に対して特に有効である。スパース性が低い場合、目的信号がアクティブである時間周波数位置 (f, τ) であっても他の妨害信号成分も多く含まれ、式(49)の近似が成り立たなくなり、角度 $\theta_{10}(f, \tau)$ が0から離れてしまうからである。すなわちスパースでない場合、例えば、図24Bのように、ある時間周波数位置 (f, τ) において目的信号 $(I(f)=1)$ とする)に対応するベクトル $V(f) \cdot H_1(f)$ の他に、妨害信号に対応するベクトル $V(f) \cdot H_2(f)$, $V(f) \cdot H_3(f)$ が共存することになる。この例の場合、白色化混合信号ベクトル $Z(f, \tau)$ は、

[0159] [数43]

$$Z(f, \tau) \approx \sum_{k=1}^3 V(f) \cdot H_k(f) \cdot S_k(f, \tau) \quad \dots (52)$$

となるため、白色化混合信号ベクトル $Z(f, \tau)$ と白色化基底ベクトル $B_{10}(f)$ とがなす角度 $\theta_{10}(f, \tau)$ も0から離れてしまう。これは目的信号がアクティブである時間周波数位置 (f, τ) の信号をマスク処理選択信号 $Y_{10}'(f, \tau)$ から排除してしまう可能性があることを示している。

また、この時間周波数マスクング処理は、目的信号のパワーが妨害信号のパワーに比べて十分に大きい場合に特に有効である。すなわち、目的信号のパワーが妨害信号のパワーに比べ十分に大きい場合、たとえスパース性が低く、目的信号がアクティブである時間周波数位置 (f, τ) に他の妨害信号成分も含まれる状況であっても、式(49)の近似が比較的成り立ち、角度 $\theta_{10}(f, \tau)$ が0に近づくからである。例えば、目的信号のパワーが妨害信号のパワーに比べ十分に大きい場合、式(52)における妨害信号の寄与度は低くなり、白色化混合信号ベクトル $Z(f, \tau)$ と白色化基底ベクトル $B_{10}(f)$ とがなす角度 $\theta_{10}(f, \tau)$ も0に近づく。これは、目的信号がアクティブである時間周波数位置 (f, τ) の信号が、マスク処理選択信号 $Y_{10}'(f, \tau)$ から排除されてしまう可能性を低くできることを示している。また、マスク処理選択信号 $Y_{10}'(f, \tau)$ に尚も残存する妨害信号成分を相対的に低くできることを意味している(ステップS105(マスクング部1152)の処理の詳細の説明終わり)。

[0160] 〔第4の実施の形態(「第2の本発明」の例)〕

次に、本発明における第4の実施の形態について説明する。

本形態は、第3の実施の形態の変形例であり、時間周波数マスクによる処理を行わない点のみが第3の実施の形態と相違する。以下では、第3の実施の形態との相違点を中心に説明し、第3の実施の形態と共通する事項については説明を省略する。

<構成>

図25は、第4の実施の形態における信号分離装置1200のブロック図の例示である。

[0161] 図25に例示するように、本形態の信号分離装置1200の信号分離装置1001との相違点は、メモリ1100に記憶領域1112, 1113を有しない点、及び時間周波数マスクング部1150を有しない点である。

<処理>

次に、本形態における信号分離装置1200の処理について説明する。

図26は、第4の実施の形態における信号分離装置1200の処理を説明するためのフローチャートである。以下では、第3の実施の形態との相違点を中心に説明する。

[0162] まず、第3の実施の形態と同様に、周波数領域変換部1120において、メモリ1100

の記憶領域1101から時間領域の混合信号 $x_q(t)$ を読み出す。そして、周波数領域変換部1120は、これらを短時間フーリエ変換などにより周波数領域の混合信号 $X_q(f, \tau)$ に変換し、メモリ1100の記憶領域1102に格納する(ステップS161)。

次に、信号分離部1130において、メモリ1100の記憶領域1102から周波数領域の混合信号 $X_q(f, \tau)$ を読み出す。この例の信号分離部1130は、読み出した混合信号 $X_q(f, \tau)$ を要素とする混合信号ベクトル $X(f, \tau) = [X_1(f, \tau), \dots, X_M(f, \tau)]^T$ に、独立成分分析(ICA)を適用し、周波数 f 毎に M 行 M 列の分離行列 $W(f) = [W_1(f), \dots, W_M(f)]^H$ ($*$ ^Hは行列 $*$ の複素共役転置行列)と分離信号ベクトル $Y(f, \tau) = W(f) \cdot X(f, \tau)$ とを算出する(ステップS162)。なお、算出された分離行列 $W(f)$ はメモリ1100の記憶領域1103に格納される。また、分離信号ベクトル $Y(f, \tau) = [Y_1(f, \tau), \dots, Y_M(f, \tau)]^T$ を構成する分離信号 $Y_p(f, \tau)$ ($p \in \{1, \dots, M\}$)は記憶領域1107に格納される。

[0163] 次に、目的信号選択部1140において、メモリ1100の記憶領域1103から分離行列 $W(f)$ を読み出し、その一般化逆行列の列である基底ベクトルを正規化し、その正規化された基底ベクトルをクラスタリングする。そして、目的信号選択部1140は、そのクラスタの分散を指標として、メモリ1100の記憶領域1107の分離信号から選択信号 $Y_{i(0)}(f, \tau)$ を周波数 f 毎に選択し、メモリ1100の記憶領域1111に格納する(ステップS163)。

次に、時間領域変換部1160において、メモリ1100の記憶領域1111から選択された分離信号 $Y_{i(0)}(f, \tau)$ を読み出し、それに短時間逆フーリエ変換などを施して時間領域の分離信号 $y(t)$ を生成し、メモリ1100の記憶領域1114に格納する(ステップS164)。

[0164] [第5の実施の形態(「第2の本発明」の例)]

次に、本発明における第5の実施の形態について説明する。

本形態は、第3の実施の形態の変形例であり、時間周波数マスクの生成方法のみが第3の実施の形態と相違する。以下では、第3の実施の形態との相違点を中心に説明し、第3の実施の形態と共通する事項については説明を省略する。

<構成>

図27は、第5の実施の形態における信号分離装置1300のブロック図の例示であ

る。また、図28Aは、図27における時間周波数マスク部1350の詳細構成を示すブロック図であり、図28Bは、図28Aのマスク生成部1351の詳細構成を示すブロック図である。なお、これらの図において第3の実施の形態と共通する部分については、第3の実施の形態の図と同じ符号を付した。

- [0165] 図27に例示するように、本形態の信号分離装置1300の信号分離装置1001との相違点は、信号分離装置1300が、時間周波数マスク部1150の代わりに時間周波数マスク部1350を有する点、メモリ1100が記憶領域1108, 1112の代わりに記憶領域1308, 1312を有する点である。また、図28Aに示すように、時間周波数マスク部1350は、マスク生成部1351及びマスク部1152を有する。また、図28Bに示すように、マスク生成部1351は、周波数正規化部1351a、ノルム正規化部1351b、セントロイド抽出部1351c、自乗距離算出部1351d及び関数生成部1351eを有している。また、周波数正規化部1351aは、第1正規化部1351aa及び第2正規化部1351abを有しており、セントロイド抽出部1351cは、セントロイド選択部1351ca及びノルム正規化部1351cbを有している。

[0166] <マスク生成処理>

本形態と第3の実施の形態との相違点は、時間周波数マスクの生成処理(ステップS104)のみである。以下では、本形態の時間周波数マスクの生成処理のみを説明する。

図29は、第5の実施の形態における時間周波数マスクの生成処理を説明するためのフローチャートである。また、図30Aは、図29におけるステップS171の詳細を説明するためのフローチャートである。また図30Bは、図29におけるステップS172の詳細を説明するためのフローチャートである。以下、これらの図を用い、本形態の時間周波数マスクの生成処理を説明する。

- [0167] まず、マスク生成部1351の周波数正規化部1351aが、メモリ1100の記憶領域1102に格納されている周波数領域の混合信号 $X(f, \tau)$ から構成される混合信号ベクトル $X(f, \tau)$ を、周波数に依存しない周波数正規化ベクトル $X'(f, \tau)$ に正規化し(周波数正規化)、当該周波数正規化ベクトル $X'(f, \tau)$ の各要素 $X'_q(f, \tau)$ をメモリ1100の記憶領域1312に格納する(ステップS171)。

[周波数正規化処理(ステップS171)の詳細]

以下に、この周波数正規化処理(ステップS171)の詳細を説明する。

[0168] まず、制御部1170(図27)においてパラメータ q に1を代入し、これを一時メモリ1180に格納する(ステップS181)。次に、周波数正規化部1351a(図28B)において、メモリ1100の記憶領域1308から前述のパラメータ d , c , Q を読み込み、記憶領域1102から各 (f, τ) に対応する混合信号ベクトル $X(f, \tau)$ の要素 $X_q(f, \tau)$ を読み込み、一時メモリ1180からパラメータ q を読み込む。そして、周波数正規化部1351aは、

[0169] [数44]

$$X_q'(f, \tau) = |X_q(f, \tau)| \exp \left[j \frac{\arg[X_q(f, \tau)/X_Q(f, \tau)]}{4fc^{-1}d} \right] \quad \dots (53)$$

の演算を行い、当該演算結果を周波数正規化ベクトル $X'(f, \tau) = [X_1'(f, \tau), \dots, X_M'(f, \tau)]^T$ の各要素として、メモリ1100の記憶領域1312に格納する(ステップS182)。なお、 $\arg[\cdot]$ は偏角を意味し、 j は虚数単位を意味する。

より詳細には、まず、周波数正規化部1351aの第1正規化部1351aaが、混合信号ベクトル $X(f, \tau)$ の特定の1つの要素 $X_q(f, \tau)$ を基準として当該混合信号ベクトル $X(f, \tau)$ の各要素 $X_q(f, \tau)$ の偏角を、以下の演算によって正規化する。

[0170] [数45]

$$X_q'''(f, \tau) = |X_q(f, \tau)| \exp \{ j \cdot \arg[X_q(f, \tau)/X_Q(f, \tau)] \} \quad \dots (54)$$

次に、周波数正規化部1351aの第2正規化部1351abが、第1正規化部1351aaで正規化された各要素 $X_q'''(f, \tau)$ の偏角を、以下のように周波数 f に比例した値 $4fc^{-1}d$ で除算する。

[0171] [数46]

$$X_q'(f, \tau) = |X_q'''(f, \tau)| \exp \left[j \frac{\arg[X_q'''(f, \tau)]}{4fc^{-1}d} \right] \quad \dots (55)$$

次に、制御部1170において、一時メモリ1180に格納されたパラメータ q が $q=M$ を満たすか否かを判断する(ステップS183)。ここで $q=M$ でなければ、制御部1170は

、 $q+1$ の演算結果を新たなパラメータ q の値とし、これを一時メモリ1180に格納し(ステップS184)、処理をステップS182へ戻す。一方、 $q=M$ であれば、制御部1170はステップS171の処理を終了させ、以下のステップS172の処理を実行させる(周波数正規化処理(ステップS171)の詳細終わり)。

- [0172] 次に、マスク生成部1351のノルム正規化部1351bが、メモリ1100の記憶領域1312に格納されている各要素 $X_q(f, \tau)$ からなる周波数正規化ベクトル $X'(f, \tau)$ を、ノルムが所定の値(この例では1)になるノルム正規化ベクトル $X''(f, \tau)$ に正規化し(ノルム正規化)、それら各要素 $X''(f, \tau)$ を記憶領域1312に格納する(ステップS172)。

[ノルム正規化処理(ステップS172)の詳細]

以下に、このノルム正規化処理(ステップS172)の詳細を説明する。

- [0173] まず、ノルム正規化部1351b(図28B)において、メモリ1100の記憶領域1312から、各 (f, τ) に対応する周波数正規化ベクトル $X'(f, \tau) = [X'_1(f, \tau), \dots, X'_M(f, \tau)]^T$ を読み込む。そして、ノルム正規化部1351bは、

- [0174] [数47]

$$\|X'(f, \tau)\| = \sqrt{\sum_{q=1}^M (X'_q(f, \tau))^2}$$

の演算を行ってそれらのノルム $\|X'(f, \tau)\|$ を求め、周波数正規化ベクトル $X'(f, \tau)$ 及びノルム $\|X'(f, \tau)\|$ を一時メモリ1180に格納する(ステップS185)。

次に、ノルム正規化部1351bは、一時メモリ1180から各 (f, τ) に対応する周波数正規化ベクトル $X'(f, \tau)$ 及びノルム $\|X'(f, \tau)\|$ を読み出し、

$$X''(f, \tau) = X'(f, \tau) / \|X'(f, \tau)\|$$

の演算を行ってノルム正規化ベクトル $X''(f, \tau)$ を求める(ステップS186)。

- [0175] 求められたノルム正規化ベクトル $X''(f, \tau)$ は、メモリ1100の記憶領域1312に格納される。これによりステップS172の処理が終了する(ノルム正規化処理(ステップS172)の詳細の説明終わり)。

次に、セントロイド抽出部1351cのセントロイド選択部1351caが、一時メモリ1180からクラスタ選択情報 l を読み込み(ステップS141参照)、このクラスタ選択情報 l

に対応するセントロイド η_k をメモリ1100の記憶領域1110から読み込む(ステップS173)。次に、ノルム正規化部1351cbが、セントロイド選択部1351caが読み込んだセントロイド η_k のノルムを所定の値(ステップS172の所定の値[この例では1])に正規化する。このノルム正規化後のセントロイド η_k をノルム正規化セントロイド η_k' と呼ぶ(ステップS174)。なお、ノルム正規化の手順は、ステップS185, S186と同様である。また、ノルム正規化セントロイド η_k' は、メモリ1100の記憶領域1312に格納される。

[0176] 次に、自乗距離算出部1351dが、メモリ1100の記憶領域1312から、ノルム正規化ベクトル $X''(f, \tau)$ とノルム正規化セントロイド η_k' とを読み込み、それらの距離の自乗

$$DS(f, \tau) = \| \eta_k' - X''(f, \tau) \|^2$$

を算出し(ステップS175)、その距離の自乗 $DS(f, \tau)$ をメモリ1100の記憶領域1312に格納する。

次に、関数生成部1351eが、メモリ1100の記憶領域1312から距離の自乗 $DS(f, \tau)$ を読み込み、これを変数とする関数を用いて時間周波数マスク $M(f, \tau)$ を生成し、メモリ1100の記憶領域1312に格納する(ステップS176)。具体的には、関数生成部1351eは、例えば、メモリ1100の記憶領域1308から実数パラメータ g, D_T を読み込み、以下のロジスティック関数を時間周波数マスク $M(DS(f, \tau))$ として生成する。なお、 D_T は事前に記憶領域1308に格納しておいたものである。また、 e はネピア数である。

[0177] [数48]

$$M(DS(f, \tau)) = \frac{1}{1 + e^{g(DS(f, \tau) - D_T)}} \quad \dots (56)$$

そして、以上のように生成された時間周波数マスク $M(DS(f, \tau))$ は、第3の実施の形態と同様、マスキング部1152でのマスキング処理に用いられる。

〔実験結果〕

第3, 第4の実施の形態の効果を示すために、マイクロホンに近い主要音声を強調して抽出する実験を行った。この実験では、図31Aに示す実験条件でインパルス応

答 $h_{qk}(r)$ を測定した。ここでは、カクテルパーティの状況を再現するようにスピーカを配置した。また、全てのスピーカの音量はほぼ等しくなるようにし、特別なスピーカの音量のみが著しく大きくならないようにした。マイクロホンでの混合は、8kHzでサンプリングした6秒間の英語音声、測定したインパルス応答と畳み込むことにより作成した。マイクロホン配置は、図31Aに示すように3次元的にしたが、信号分離装置を実装したシステム(装置)には、基準マイクロホン(Mic.2)とその他のマイクロホン間の最大距離(3.5cm)のみを情報として与え、詳細なマイクロホン配置は与えていない。各実験において、マイクロホンに近い4つのスピーカ位置(a120,b120,c120,c170)の中から1つを目的音源として選び、残りの3つのスピーカは無音状態にした。マイクロホンから離れた6個のスピーカは、各実験で常に妨害音として作動させた。抽出結果はInput SIR—Output SIRであるSIR(signal—to—interference ratio)の改善量によって評価した。この値が大きいほど、より良く目的音声を抽出し、他の妨害音をより抑圧できたことを意味する。これら2種類のSIRは、

[0178] [数49]

$$\text{InputSIR} = 10 \log_{10} \frac{\langle |\sum_r h_{11}(r) \cdot s_1(t-r)|^2 \rangle_t}{\langle |\sum_{k \neq 1} \sum_r h_{1k}(r) \cdot s_k(t-r)|^2 \rangle_t} \quad (\text{dB})$$

$$\text{OutputSIR} = 10 \log_{10} \frac{\langle |\sum_r u_{11}(r) \cdot s_1(t-r)|^2 \rangle_t}{\langle |\sum_{k \neq 1} \sum_r u_{1k}(r) \cdot s_k(t-r)|^2 \rangle_t} \quad (\text{dB})$$

によって定義される。ここで、

$$u_{1k}(r) = \sum_{q=1}^M \sum_{\tau=0}^{L-1} w_{1q}(\tau) \cdot h_{qk}(r-\tau)$$

は $s_k(t)$ から $y_1(t)$ へのインパルス応答である。

実験は、各目的音源位置に対して、7つの音声(目的音声1個と妨害音声6個)からなる組合せを16種類作成して行った。図31Bは、ICAのみの場合(第4の実施の形態)と、ICAと時間周波数マスキングの両方を組み合わせた場合(第3の実施の形態)とのSIRの平均改善量を示す表である。SIRの改善量は、目的音源の位置によって多少左右されるが、おおむね良好な結果を得ている。位置a120とb120とでの結果が良

いのは、妨害音が異なった位置から到来するからである。位置c120とc170は、2次元的に見ると、同様の方向から多くの妨害音が到来するため困難な位置に思えるが、実際に位置c170での結果は大変良かった。これは、c170の高さが妨害音と異なり、3次元状に配列されたマイクロホンによって、本システムが高さの違いを自動的に利用して抽出処理を行っているからである。また、図31Bの表によると、時間周波数マスキングにより性能が改善されることが分かる。時間周波数マスクを決定する式(46)に対して図31Aに示す3種類のパラメータを用いた。より小さい θ_T を用いることで、より大きいSIR改善量が達成されている。しかしながら、小さい θ_T によるいくつかの音は、不自然な音(musical noise)を伴った。実験の結果、多くの場合においてパラメータ(θ_T, g)=($0.333\pi, 20$)が、十分に妨害音の抑圧を行いながら、かつ自然な音を出力することを確認した。

[0179] 〔変形例等〕

なお、本発明は上述の第3から第5の実施の形態に限定されるものではない。例えば、上述した各実施の形態では、信号分離部1130においてM行M列の分離行列 $W(f)$ を算出することとしたが、N行M列など正方行列以外の分離行列 $W(f)$ を算出する構成としてもよい。この場合、基底ベクトルは、分離行列 $W(f)$ の一般化逆行列 $W^+(f)$ (例えばムーア・ペンローズ(Moore-Penrose)型一般化逆行列)の各列となる。

また、第3の実施の形態では、時間周波数マスクを用いて選択信号 $Y_{10}(f, \tau)$ から妨害信号成分をさらに抑制したマスク処理選択信号 $Y_{10}'(f, \tau)$ を生成することとしたが、他の方法によって妨害信号成分を抑制したマスク処理選択信号 $Y_{10}'(f, \tau)$ を生成することとしてもよい。例えば、信号源が2個だけである場合、抽出された分離信号 $Y_1(f, \tau), Y_2(f, \tau)$ の大きさを比較し、 $|Y_1(f, \tau)| > |Y_2(f, \tau)|$ ならば $Y_1(f, \tau)$ をマスク処理選択信号 $Y_{10}'(f, \tau)$ として抽出し、 $|Y_1(f, \tau)| < |Y_2(f, \tau)|$ ならば $Y_2(f, \tau)$ をマスク処理選択信号 $Y_{10}'(f, \tau)$ として抽出する時間周波数マスクを生成し、分離信号 $Y_1(f, \tau), Y_2(f, \tau)$ を要素とするベクトルに掛け合わせることもよい。

[0180] さらに、上述の第3の実施の形態では、信号分離部1130が独立成分分析(ICA)を用いて分離行列と分離信号とを算出することとしたが、時間周波数マスク(時間周波数毎のマスクを意味し、例えば、1か0の値をとるバイナリマスク)によって観測信号か

ら分離信号を抽出し(例えば、「O. Yilmaz and S. Rickard, "Blind separation of speech mixtures via time-frequency masking," IEEE Trans. on SP, vol. 52, no. 7, pp. 1830-1847, 2004.」参照)、その算出結果から分離行列を生成することとしてもよい。

また、上述の第3の実施の形態では、周波数正規化部1142aの第1正規化部1142aaが、基底ベクトル $A_p(f)$ の特定の1つの要素 $A_{qp}(f)$ を基準とし、式(35)の一部である式(15)によって当該基底ベクトル $A_p(f)$ の各要素 $A_{qp}(f)$ の偏角を正規化することとした。しかし、例えば、前述した式(27-1)～(27-3)等に従って、第1正規化部1142aaが、基底ベクトル $A_p(f)$ の特定の1つの要素 $A_{qp}(f)$ を基準として、当該基底ベクトル $A_p(f)$ の各要素 $A_{qp}(f)$ の偏角を正規化する構成であってもよい。

[0181] また、周波数正規化部1142aが、式(35)の代わりに、前述した式(28-1)～(28-4)等の演算によって周波数正規化を行うこととしてもよい。

また、上述の第3の実施の形態では、ノルム正規化部1142bにおいてノルムが1になるような正規化を行ったが、ノルムが1以外の規定値となるように正規化を行ってもよい。さらに、ノルム正規化部1142bを設けず、ノルム正規化を行わない構成であってもよい。この場合のクラスタリングは、前述した通り、ベクトルの方向の類似度を基準にして行う。

[0182] また、パラメータdは、全てのセンサqに対して同一であってもよいし、各センサqに対応して複数の値を設定しておいてもよい。例えば、基準センサとセンサqとの距離を、それぞれのセンサqに対応するパラメータdの値としてもよい。

〔第6の実施の形態(「第3の本発明」の例)〕

次に、本発明における第6の実施の形態について説明する。

本形態は、前述の原理を用い、センサの厳密な配置情報を必要とすることなく、全ての観測信号から得られた情報を容易かつ効率的に利用して信号分離を行う形態である。本形態では、後述する「混合信号ベクトル」が前述の「複素ベクトル」に相当する。

[0183] <構成>

本形態の信号分離装置2001も、第1の実施の形態と同様、公知のノイマン型のコンピュータに信号分離プログラムが読み込まれることにより構成される。図32は、第6

の実施の形態における信号分離装置2001の構成を例示したブロック図である。また、図33は、図32における信号分離部2120の詳細を例示したブロック図である。なお、これらの図における実線の矢印は実際のデータの流れを示し、破線の矢印は理論的な情報の流れを示す。また、これらの図において制御部2140に出入りするデータの流れに対応する矢印は省略してある。

[0184] 図32、図33に例示するように、信号分離装置2001は、メモリ2100、周波数領域変換部2110(「複素ベクトル生成部」に相当する機能を含む)、信号分離部2120、時間領域変換部2130及び制御部2140を有している。また、この信号分離部2120は、周波数正規化部2121(「正規化部」を構成)、ノルム正規化部2122(「正規化部」を構成)、クラスタリング部2123及び分離信号生成部2124を有している。また、周波数正規化部2121は、第1正規化部2121a及び第2正規化部2121bを有している。また、制御部2140は一時メモリ2141を有している。

[0185] ここでメモリ2100及び一時メモリ2141は、レジスタ10ac、補助記憶装置10f及びRAM10d等に相当する。また、周波数領域変換部2110、信号分離部2120、時間領域変換部2130及び制御部2140は、CPU10aにOSプログラムや信号分離プログラムが読み込まれ、CPU10aがこれらを実行することにより構成されるものである。

<処理>

次に、本形態の信号分離装置2001の処理について説明する。なお、以下では、N個の源信号が混合され、M個のセンサで観測された状況を取り扱う。また、前処理において、各センサで観測された時間領域の混合信号 $x_q(t)$ ($q=1, \dots, M$) がメモリ2100の記憶領域2101に格納され、信号の伝達速度 c 、M以下の自然数から選択された基準値 Q, Q' (それぞれM個のセンサから選択された基準センサの添字番号) 及び実数 d の各パラメータが記憶領域2105に格納されているものとする。

[0186] 図34は、本形態における信号分離装置2001の処理の全体を説明するためのフローチャートである。以下、この図に沿って、本形態における信号分離装置2001の処理を説明していく。

[処理の全体]

まず、周波数領域変換部2110において、メモリ2100の記憶領域2101から時間領

域の混合信号 $x_q(t)$ を読み出し、これらを短時間離散フーリエ変換等によって周波数毎の時系列信号(「周波数領域の混合信号」と呼ぶ) $X_q(f, \tau)$ ($q=1, \dots, M, f=0, f_s/L, \dots, f_s(L-1)/L, f_s$ はサンプリング周波数)に変換し、メモリ2100の記憶領域2102に格納する(ステップS201)。

[0187] 次に、信号分離部2120の周波数正規化部2121において、メモリ2100の記憶領域2102から周波数領域の混合信号 $X_q(f, \tau)$ を読み出す。周波数領域の混合信号 $X_q(f, \tau)$ を読み出した周波数正規化部2121は、これらからなる混合信号ベクトル $X(f, \tau)=[X_1(f, \tau), \dots, X_M(f, \tau)]^T$ を、周波数 f に依存しない周波数正規化ベクトル $X'(f, \tau)$ に正規化する(ステップS202)。生成された各周波数正規化ベクトル $X'(f, \tau)$ はメモリ2100の記憶領域2103に格納される。なお、ステップS202の処理の詳細については後述する。

[0188] 次に、信号分離部2120のノルム正規化部2122において、メモリ2100の記憶領域2103から各周波数正規化ベクトル $X'(f, \tau)$ を読み込み、これらをノルムが所定の値(例えば1)になるノルム正規化ベクトル $X''(f, \tau)$ に正規化する。そして、ノルム正規化部2122は、生成した各ノルム正規化ベクトル $X''(f, \tau)$ をメモリ2100の記憶領域2104に格納する(ステップS203)。なお、この詳細については後述する。

次に、信号分離部2120のクラスタリング部2123において、メモリ2100の記憶領域2104から各ノルム正規化ベクトル $X''(f, \tau)$ を読み込み、これらをクラスタリングし、クラスタを生成する。そして、クラスタリング部2123は、各クラスタを特定するクラスタ情報 C_k ($k=1, \dots, N$) 番目のクラスタのメンバ $X''(f, \tau)$ を特定する情報)をメモリ2100の記憶領域2106に格納する(ステップS204)。なお、この詳細については後述する。

[0189] 次に、信号分離部2120の分離信号生成部2124において、メモリ2100の記憶領域2106, 2105からクラスタ情報 C_k 及び基準値 Q' を読み込む。そして、分離信号生成部2124は、クラスタ情報 C_k 及び基準値 Q' を用い、 k 番目のクラスタに属するノルム正規化ベクトル $X''(f, \tau)$ に対応する混合信号ベクトル $X(f, \tau)$ から Q' 番目の要素 $X_{Q'}(f, \tau)$ を記憶領域2102から抽出し、これを k 番目の要素 $Y_k(f, \tau)$ とした分離信号ベクトル $Y(f, \tau)$ を生成する。そして、分離信号生成部2124は、生成した分離信号ベクトル $Y(f, \tau)$ をメモリ2100の記憶領域2107に格納する(ステップS205)。なお、この

詳細については後述する。

[0190] 最後に、時間領域変換部2130において、メモリ2100の記憶領域2107から分離信号ベクトル $Y(f, \tau)$ を読み込み、その分離信号成分 $Y_k(f, \tau)$ を添字 k 毎に短時間逆フーリエ変換等により、時間領域の分離信号 $y_k(t)$ に変換する。そして、時間領域変換部2130は、変換した時間領域の分離信号 $y_k(t)$ をメモリ2100の記憶領域2108に格納する(ステップS206)。

次に、各処理の詳細について説明する。

[周波数正規化部2121・ノルム正規化部2122の処理の詳細]

周波数正規化部2121及びノルム正規化部2122は、全ての混合信号ベクトル $X(f, \tau) = [X_1(f, \tau), \dots, X_M(f, \tau)]^T$ ($f=0, f/L, \dots, f(L-1)/L$)を正規化し、それらを周波数に依存せず、信号源の位置のみに依存するノルム正規化ベクトル $X''(f, \tau)$ に正規化する。これにより、それらをステップS204でクラスタリングした際に、各クラスタが各信号源のみに対応するようになる。この正規化を適切に行わないとクラスタが形成されない。前述した通り、本形態における正規化は、周波数正規化とノルム正規化とからなる。周波数正規化は、周波数正規化部2121において、混合信号ベクトル $X(f, \tau)$ を周波数に依存しない周波数正規化ベクトル $X'(f, \tau)$ に正規化するものである。ノルム正規化は、ノルム正規化部2122において、周波数正規化ベクトル $X'(f, \tau)$ をノルムが規定値(この例では1)をとるノルム正規化ベクトル $X''(f, \tau)$ に正規化するものである。以下にこれらの正規化の詳細を説明する。

[0191] [周波数正規化部2121における処理の詳細(ステップS202の処理の詳細)]

図35Aは、図34に示したステップS202の処理の詳細を説明するためのフローチャートである。以下、この図に沿ってステップS202の処理の詳細を説明する。

まず、制御部2140(図32)においてパラメータ q に1を代入し、これを一時メモリ2141に格納する(ステップS211)。次に、周波数正規化部2121(図32, 図33)において、メモリ2100の記憶領域2105から前述のパラメータ d, c, Q を読み込み、記憶領域2102から各 (f, τ) に対応する混合信号ベクトル $X(f, \tau)$ の要素 $X_q(f, \tau)$ を読み込み、一時メモリ2141からパラメータ q を読み込む。そして、周波数正規化部2121は、

[0192] [数50]

$$X_q'(f, \tau) = |X_q(f, \tau)| \exp \left[j \frac{\arg[X_q(f, \tau)/X_Q(f, \tau)]}{4fc^{-1}d} \right] \dots (60)$$

の演算を行い、当該演算結果を周波数正規化ベクトル $X'(f, \tau) = [X_1'(f, \tau), \dots, X_M'(f, \tau)]^T$ の各要素として、メモリ2100の記憶領域2103に格納する(ステップS212)。なお、 $\arg[\cdot]$ は偏角を意味し、 j は虚数単位を意味する。

より詳細には、まず、周波数正規化部2121の第1正規化部2121aが、混合信号ベクトル $X(f, \tau)$ の特定の1つの要素 $X_q(f, \tau)$ を基準として当該混合信号ベクトル $X(f, \tau)$ の各要素 $X_q(f, \tau)$ の偏角を、以下の演算によって正規化する。

[0193] [数51]

$$X_q'''(f, \tau) = |X_q(f, \tau)| \exp \{ j \cdot \arg[X_q(f, \tau)/X_Q(f, \tau)] \} \dots (61)$$

次に、周波数正規化部2121の第2正規化部2121bが、第1正規化部2121aで正規化された各要素 $X_q'''(f, \tau)$ の偏角を、以下のように周波数 f に比例した値 $4fc^{-1}d$ で除算する。

[0194] [数52]

$$X_q'(f, \tau) = |X_q'''(f, \tau)| \exp \left[j \frac{\arg[X_q'''(f, \tau)]}{4fc^{-1}d} \right] \dots (62)$$

次に、制御部2140において、一時メモリ2141に格納されたパラメータ q が $q=M$ を満たすか否かを判断する(ステップS213)。ここで $q=M$ でなければ、制御部2140は、 $q+1$ の演算結果を新たなパラメータ q の値とし、これを一時メモリ2141に格納し(ステップS214)、処理をステップS212へ戻す。一方、 $q=M$ であれば、制御部2140はステップS202の処理を終了させ、以下のステップS203の処理を実行させる。

[0195] [ノルム正規化部2122における処理の詳細(ステップS203の詳細)]

図35Bは、図34に示したステップS203の処理の詳細を説明するためのフローチャートである。以下、この図に沿ってステップS303の処理の詳細を説明する。

ノルム正規化部2122(図32、図33)において、メモリ2100の記憶領域2103から、各 (f, τ) に対応する周波数正規化ベクトル $X'(f, \tau) = [X_1'(f, \tau), \dots, X_M'(f, \tau)]^T$ を読み

込む。そして、ノルム正規化部2122は、

[0196] [数53]

$$\|X'(f, \tau)\| = \sqrt{\sum_{q=1}^M (X_q'(f, \tau))^2}$$

の演算を行ってそれらのノルム $\|X'(f, \tau)\|$ を求め、周波数正規化ベクトル $X'(f, \tau)$ 及びノルム $\|X'(f, \tau)\|$ を一時メモリ2141に格納する(ステップS221)。

次に、ノルム正規化部2122は、一時メモリ2141から各 (f, τ) に対応する周波数正規化ベクトル $X'(f, \tau)$ 及びノルム $\|X'(f, \tau)\|$ を読み出し、

$$X''(f, \tau) = X'(f, \tau) / \|X'(f, \tau)\| \quad \dots (63)$$

の演算を行ってノルム正規化ベクトル $X''(f, \tau)$ を求める(ステップS222)。求められた各ノルム正規化ベクトル $X''(f, \tau)$ は、メモリ2100の記憶領域2104に格納される。これによりステップS203の処理が終了する。

[0197] このように生成されたノルム正規化ベクトル $X''(f, \tau)$ は、周波数に依存せず、信号源の位置のみに依存するベクトルとなる。その結果、このノルム正規化ベクトル $X''(f, \tau)$ はクラスタを形成することになる。以下にこの理由を説明する。

[ノルム正規化ベクトル $X''(f, \tau)$ がクラスタを形成する理由]

本形態では源信号のスパース性を仮定しているので、混合信号ベクトル $X(f, \tau)$ の各要素 $X_q(f, \tau)$ は、源信号 p に対応する信号源 k からセンサ q への周波数応答 H_{qk} に比例した(ある複素数スカラーである源信号 $S_k(f, \tau)$ が掛かった)ものになっている($X_q(f, \tau) = H_{qk}(f, \tau) \cdot S_k(f, \tau)$)。

[0198] これらの源信号 $S_k(f, \tau)$ は離散時間に応じて(すなわち位相に応じて)変化するが、当然ながら、周波数 f が同じであれば、センサ q で観測される源信号 $S_k(f, \tau)$ の偏角と、基準センサ Q で観測される源信号 $S_k(f, \tau)$ の偏角との相対値は離散時間が変化しても一定である。

前述のように、周波数正規化部2121の第1正規化部2121aは、混合信号ベクトル $X(f, \tau)$ の特定の1つの要素 $X_q(f, \tau)$ を基準として当該混合信号ベクトル $X(f, \tau)$ の各 $X_q(f, \tau)$ の偏角を正規化する。

[0199] これにより、上述の源信号 $S_k(f, \tau)$ の位相に起因する不確定性を取り除き、源信号 p

とセンサ q とに対応する混合信号ベクトル $X(f, \tau)$ の各要素 $X_q(f, \tau)$ の偏角を、源信号 p と基準センサ Q (基準値 Q に対応)とに対応する混合信号ベクトル $X(f, \tau)$ の要素 $X_Q(f, \tau)$ の偏角に対する相対値として表現する。なお、この場合、要素 $X_q(f, \tau)$ の偏角に対応する相対値は0と表現される。

信号源 k からセンサ q への周波数応答を、反射や残響の無い直接波モデルで近似して考える。すると、上記の第1正規化部2121aにより正規化された偏角は、信号源 k からセンサへの波の到達時間差と周波数 f の双方に比例したものになる。ここでの到達時間差とは、信号源 k からの波がセンサ q に到達するまで時間と、当該波が基準センサ Q に到達するまでの時間との時間差である。

[0200] また、前述のように、第2正規化部2121bは、第1正規化部2121aで正規化された各要素 $X_q''(f, \tau)$ の偏角を周波数 f に比例した値で除算する。これにより、各要素 $X_q''(f, \tau)$ を、それらの偏角の周波数依存性を排除した各要素 $X_q'(f, \tau)$ に正規化する。これにより、正規化された各要素 $X_q'(f, \tau)$ は、直接波モデルに従えば、信号源 k からセンサへの波の到達時間差のみに依存するものとなる。ここで、信号源 k からセンサへの波の到達時間差は、信号源 k 、センサ q 、基準センサ Q の相対位置にのみ依存する。そのため、信号源 k 、センサ q 、基準センサ Q が同じであれば、周波数 f が異なっても各要素 $X_q'(f, \tau)$ の偏角は同一となる。従って、周波数正規化ベクトル $X'(f, \tau)$ は、周波数 f には依存せず、信号源 k の位置のみに依存する。そのため、周波数正規化ベクトル $X'(f, \tau)$ のノルムを正規化したノルム正規化ベクトル $X''(f, \tau)$ のクラスタリングによって、同じ信号源毎に対応するクラスタが形成される。なお、実際の環境では、反射や残響などの影響により、直接波モデルは厳密には満たされないが、後述の実験結果に示すとおり十分に良い近似となっている。

[0201] 次に、ノルム正規化ベクトル $X''(f, \tau)$ がクラスタを形成する理由を、モデルを用いて説明する。

前述した式(1)に示されるインパルス応答 $h_{qk}(r)$ を直接波(近距離場)混合モデルを用いて近似し、周波数領域で表すと、

[0202] [数54]

$$H_{qk}(f) = \frac{\gamma(f)}{d_{qk}} \exp[-j2\pi f c^{-1}(d_{qk} - d_{Qk})] \quad \dots (64)$$

となる。ここで、 d_{qk} は信号源 k とセンサ q との間の距離であり、 $\gamma(f)$ は周波数に依存した定数である。また、減衰 $\gamma(f)/d_{qk}$ は、距離 d_{qk} と定数 $\gamma(f)$ によって決まり、遅延 $(d_{qk} - d_{Qk})/c$ は、基準センサ Q の位置で正規化された距離によって決まる。

また、信号のスパース性を仮定すると、各時間周波数 (f, τ) において以下の関係が成り立つ。

$$[0203] \quad X_q(f, \tau) = H_{qk}(f, \tau) \cdot S_k(f, \tau) \quad \dots (65)$$

ここで、式(62) (63) (64) (65)より、

[0204] [数55]

$$X_p''(f, \tau) = \frac{1}{d_{qk} D} \exp\left[-j\frac{\pi}{2} \frac{(d_{qk} - d_{Qk})}{d}\right], \quad D = \sqrt{\sum_{i=1}^M \frac{1}{d_{ik}^2}} \quad \dots (66)$$

となる。

この式から分かるように、ノルム正規化ベクトル $X''(f, \tau)$ の各要素 $X_q''(f, \tau)$ は、周波数 f からは独立であり、信号源 k とセンサ q の位置のみに依存する。従って、ノルム正規化ベクトルをクラスタリングすると、同じ信号源ごとに対応するクラスが形成される。

[0205] また、同様なことは、信号の減衰を考慮していない近距離場混合モデルや遠距離場混合モデルでモデル化した場合にも言える(1の実施の形態と同様)。

また、第1の実施の形態と同様、式(66)から、パラメータ d の値は、 $d > d_{\max}/2$ であることが望ましく(d_{\max} は要素 $X_q''(f, \tau)$ に対応する基準センサと他のセンサとの最大距離を意味する。)、より好ましくは $d \geq d_{\max}$ であることが望ましく、さらにより好ましくは $d = d_{\max}$ であることが望ましいことが分かる。

図37及び図38は、パラメータ d 毎のノルム正規化ベクトル $X''(f, \tau)$ の要素 $X_q''(f, \tau)$ と、その偏角 $\arg[X_q''(f, \tau)]$ との関係を説明するための複素平面図である。なお、これらにおける横軸は実軸を縦軸は虚軸を示している。

[0206] 図37Aは $d_{\max}/2 \leq d$ の場合における複素平面図である。ここで上述の d_{\max} の定義よ

り、任意の q 及び k に対し $d_{qk} - d_{Qk}$ の絶対値は d_{\max} 以下となる。よって $d_{\max}/2 \geq d$ の場合、 $(\pi/2) \cdot (d_{qk} - d_{Qk})/d \leq -\pi$, $(\pi/2) \cdot (d_{qk} - d_{Qk})/d \geq \pi$ と成り得る。その結果、式(66)で表される $X_q''(f, \tau)$ の偏角 $\arg[X_q''(f, \tau)]$ は、 2π を超える $\alpha_1 \leq \arg[X_q''(f, \tau)] \leq \alpha_2$ ($\alpha_1 \leq -\pi$, $\alpha_2 \geq \pi$)の範囲に分布する可能性がある。そのため、異なるノルム正規化ベクトル $X_q''(f, \tau)$ の要素 $X_q''(f, \tau)$ の偏角が一致する可能性があり、前述したクラスタリングにおいて、異なるノルム正規化ベクトル $X_q''(f, \tau)$ を同じクラスにクラスタリングしてしまう可能性がある。よって、 $d > d_{\max}/2$ であることが望ましい。しかし、この偏角の重複範囲に対応するノルム正規化ベクトル $X_q''(f, \tau)$ のサンプルが存在しないならば $d_{\max}/2 \geq d$ としても問題はない。

[0207] 図37Bは $d_{\max}/2 < d < d_{\max}$ の場合における複素平面図である。この場合、 $-\pi < (\pi/2) \cdot (d_{qk} - d_{Qk})/d < -\pi/2$, $\pi/2 < (\pi/2) \cdot (d_{qk} - d_{Qk})/d < \pi$ と成り得る。その結果、式(66)で表される $X_q''(f, \tau)$ の偏角 $\arg[X_q''(f, \tau)]$ は、 $\beta_1 \leq \arg[X_q''(f, \tau)] \leq \beta_2$ ($-\pi < \beta_1 < -\pi/2$, $\pi/2 < \beta_2 < \pi$)の範囲に分布する可能性がある。そのため、 $-\pi < \arg[X_q''(f, \tau)] < -\pi/2$ 及び $\pi/2 < \arg[X_q''(f, \tau)] < \pi$ の範囲において、異なるノルム正規化ベクトル $X_q''(f, \tau)$ の要素間における偏角の差の増加に伴い、これらの要素間の距離が単調増加しないこともありうる。これは、前述したクラスタリングの精度を低下させる可能性がある。よって $d \geq d_{\max}$ であることがより望ましい。

[0208] 図38Aは $d = d_{\max}$ の場合における複素平面図であり、図38Bは $d > d_{\max}$ の場合における複素平面図である。ここで $d > d_{\max}$ の場合、 $-\pi/2 < (\pi/2) \cdot (d_{qk} - d_{Qk})/d < 0$, $0 < (\pi/2) \cdot (d_{qk} - d_{Qk})/d < \pi/2$ と成り得る。その結果、式(66)で表される $X_q''(f, \tau)$ の偏角 $\arg[X_q''(f, \tau)]$ は、図38Bに示すように、 $\gamma_1 \leq \arg[X_q''(f, \tau)] \leq \gamma_2$ ($-\pi/2 < \gamma_1 < 0$, $0 < \gamma_2 < \pi/2$)の範囲に分布する。そして、 d が大きくなればなるほどその分布範囲は狭くなっていき、狭い範囲にクラスが密集していく。これは、前述したクラスタリングの精度を低下させる。

[0209] これに対し、 $d = d_{\max}$ である場合、 $-\pi/2 \leq (\pi/2) \cdot (d_{qk} - d_{Qk})/d < 0$, $0 < (\pi/2) \cdot (d_{qk} - d_{Qk})/d \leq \pi/2$ と成り得る。その結果、式(66)で表される $X_q''(f, \tau)$ の偏角 $\arg[X_q''(f, \tau)]$ は図38Aに示すように $-\pi/2 \leq \arg[X_q''(f, \tau)] \leq \pi/2$ の範囲に分布する。この場合、ノルム正規化ベクトル $X_q''(f, \tau)$ の要素間における偏角の差の増加に対して、

それらの距離も単調増加するという関係を維持しつつ、できるだけ広い範囲にクラスタを分散させることができる。その結果、一般的にクラスタリングの精度を向上させることができる([周波数正規化部2121・ノルム正規化部2122の処理の詳細]の説明終わり)。

[0210] [クラスタリング部2123の処理の詳細(ステップS204の詳細)]

前述のようにクラスタリング部2123は、メモリ2100の記憶領域2104からノルム正規化ベクトル $X''(f, \tau)$ を読み込み、これらをクラスタリングしてM個のクラスタを生成する。このクラスタリングは、例えば、各クラスタのメンバ($X''(f, \tau) \in C_k$)と各クラスタのセントロイド η_k との間の二乗和 U_k の総和 U

[0211] [数56]

$$U = \sum_{k=1}^M U_k$$

$$U_k = \sum_{X''(f, \tau) \in C_k} \|X''(f, \tau) - \eta_k\|^2$$

を最小化することを基準に行われる。この最小化は、例えば、非特許文献6などで解説されているk-meansクラスタリングを用いることによって効果的に行うことができる。なお、クラスタ情報 C_k が示すクラスタのセントロイド(中心ベクトル) η_k は、

[0212] [数57]

$$\eta_k = \frac{\sum_{X''(f, \tau) \in C_k} X''(f, \tau) / |C_k|}{\left\| \sum_{X''(f, \tau) \in C_k} X''(f, \tau) / |C_k| \right\|}$$

によって計算される。ここで $|C_k|$ はクラスタ情報 C_k が示すクラスタのメンバ(ノルム正規化ベクトル $X''(f, \tau)$)の数である。また、ここでは、距離としてユークリッド距離の自乗を用いているが、これを一般化したミンコフスキー距離などを用いてもよい([クラスタリング部2123の処理の詳細]の説明終わり)。

[分離信号生成部2124の処理の詳細(ステップS205の詳細)]

図36は、図34に示したステップS205の処理の詳細を説明するためのフローチャートである。以下、この図に沿ってステップS205の処理の詳細を説明する。

[0213] まず、制御部2140(図32)において、全てのパラメータ $k(k=1, \dots, N)$ 及び時間

周波数 (f, τ) (定義された範囲における全ての f と τ)に対する $Y_k(f, \tau)$ の値を0に初期化し、これらをメモリ2100の記憶領域2107に格納する(ステップS230)。

次に、制御部2140においてパラメータ k に1を代入し、これを一時メモリ2141に格納する(ステップS231)。次に、分離信号生成部2124(図32, 図33)において、メモリ2100の記憶領域2106のクラスタ情報 C_k を読み込み、これが示す k 番目のクラスタのメンバ(ノルム正規化ベクトル $X''(f, \tau)$)を抽出し、これらを一時メモリ2141に格納する(ステップS232)。次に、分離信号生成部2124は、ステップS232で一時メモリ2141に格納された各ノルム正規化ベクトル $X''(f, \tau)$ を参照し、これらに対応する時間周波数 (f, τ) の各混合信号ベクトル $X(f, \tau)$ をメモリ2100の記憶領域2102から読み込み、これらを一時メモリ2141に格納する(ステップS233)。次に、分離信号生成部2124は、メモリ2100の記憶領域2105から基準値 Q' を読み込み、一時メモリ2141から、ステップS233で格納された混合信号ベクトル $X(f, \tau)$ の Q' 番目の要素 $X_{Q'}(f, \tau)$ を抽出する(各時間周波数 (f, τ) について)。そして、分離信号生成部2124は、抽出した要素 $X_{Q'}(f, \tau)$ を分離信号ベクトル $Y(f, \tau)$ の k 番目の要素 $Y_k(f, \tau)$ としてメモリ2100の記憶領域2107の値を更新する(ステップS234)。すなわち、この例の分離信号生成部2124は、

[0214] [数58]

$$Y_k(f, \tau) = \begin{cases} X_{Q'}(f, \tau) & X''(f, \tau) \in C_k \\ 0 & \text{otherwise} \end{cases}$$

として要素 $Y_k(f, \tau)$ を抽出する。

次に、制御部2140において、一時メモリ2141に格納されたパラメータ k が $k=N$ を満たすか否かを判断する(ステップS235)。ここで $k=N$ でなければ、制御部2140は、 $k+1$ の演算結果を新たなパラメータ k の値とし、これを一時メモリ2141に格納し(ステップS236)、処理をステップS232へ戻す。一方、 $k=N$ であれば、制御部2140はステップS205の処理を終了させる([分離信号生成部2124の処理の詳細]の説明終わり)。

[0215] <実験結果>

次に、本形態による音源分離実験結果を示す。ここでは、本形態の効果を示すた

めに二種類の信号分離の実験を行った。

1つ目の実験は、2つのセンサによる分離実験である。その実験条件は図39Aに示す通りである。信号源の数は3つであり、6秒間の英語の音声をスピーカから流した。また、この実験結果を図39Bの表に示す。この表では、SIR(signal-to-interference ratio)の向上量を示している。大きい数字ほど分離性能が良いことを示す。ここでは、分離前のセンサでの観測結果[InputSIR]と、DOAによるクラスタリングによる結果[DOA(Previous)]と、本形態(正規化によるクラスタリング)による結果[Normalized observer vector(Proposed)]を示している。この結果より、2つのセンサの場合、本形態の方法によって、DOAのみによるクラスタリング結果と同等の性能で信号分離が可能であることが確認できた。

- [0216] 2つ目の実験は、不規則なセンサ配置によるものである。実験条件は図40Aに示す通りである。この実験では、4つの無指向性マイク(センサ)を直線的ではない配置とした。また、これらの配置情報としては、マイク間隔の上限が4cmであるということだけを分離システムに与えた。信号源の数は4つであり、6秒間の英語の音声をスピーカから流した。このようなセンサや信号源の配置においてDOAを使おうとすると、それぞれのセンサペア毎にDOAを推定→センサペア毎にクラスタリング→全てのセンサペアについてのクラスタリング結果を統合、という複雑な処理を行う必要がある。しかし、本形態の方法ではそのような複雑な統合処理を行うことなく、図40Bの表に示すような高い分離性能を得ることができる。また、2つ目の実験を図41Aのような条件で行った場合も同様に図41Bの表に示すような高い分離性能を得ることができた。

[0217] <第6の実施の形態の特徴>

以上より、上述した実施の形態の特徴をまとめると以下のようになる。

(1)混合信号ベクトルから得られる情報をすべて用いてクラスタリングを行うため、全てのセンサの情報を有効に活用でき、信号分離の性能が向上する。

(2)センサの配置情報を厳密に知る必要がないため、不規則なセンサ配置を採用でき、さらにセンサ位置をキャリブレーションする必要が無い。

<変形例等>

なお、本発明は、上述の第6の実施の形態に限定されるものではない。例えば、第

6の実施の形態では、周波数正規化部2121の第1正規化部2121aが、混合信号ベクトル $X(f, \tau)$ の特定の1つの要素 $X_q(f, \tau)$ を基準として当該混合信号ベクトル $X(f, \tau)$ の各要素 $X_q(f, \tau)$ の偏角を、式(61)の演算によって正規化することとした。しかし、例えば、以下のような式に従って、周波数正規化部2121の第1正規化部2121aが、混合信号ベクトル $X(f, \tau)$ の特定の1つの要素 $X_q(f, \tau)$ を基準として当該混合信号ベクトル $X(f, \tau)$ の各要素 $X_q(f, \tau)$ の偏角を正規化する構成であってもよい。

[0218] [数59]

$$\begin{aligned} X_q'''(f, \tau) &= |X_q(f, \tau)| \exp\{j \cdot (\arg[X_q(f, \tau) \cdot X_Q^*(f, \tau)])\} \\ X_q'''(f, \tau) &= |X_q(f, \tau)| \exp\{j \cdot (\arg[X_q(f, \tau)] - \arg[X_Q(f, \tau)])\} \\ X_q'''(f, \tau) &= |X_q(f, \tau)| \exp\{j \cdot \phi(\arg[X_q(f, \tau)/X_Q(f, \tau)])\} \end{aligned}$$

ただし、 \cdot^* は、 \cdot の複素共役である。また、 $\phi\{\cdot\}$ は関数であり、クラスタリング精度の観点から好ましくは単調増加関数であることが望ましい。

また、周波数正規化部2121が式(60)の代わりに、

[0219] [数60]

$$\begin{aligned} X_q'(f, \tau) &= \rho \cdot \frac{\arg[X_q(f, \tau)/X_Q(f, \tau)]}{4fc^{-1}d} \\ X_q'(f, \tau) &= \rho \cdot \frac{\arg[X_q(f, \tau) \cdot X_Q^*(f, \tau)]}{4fc^{-1}d} \\ X_q'(f, \tau) &= \rho \cdot \frac{\arg[X_q(f, \tau)] - \arg[X_Q(f, \tau)]}{4fc^{-1}d} \\ X_q'(f, \tau) &= \rho \cdot \frac{\phi(\arg[X_q(f, \tau)/X_Q(f, \tau)])}{4fc^{-1}d} \end{aligned}$$

等の演算によって周波数正規化を行うこととしてもよい。ただし、 ρ は定数である(例えば、 $\rho = 1$)。

また、上述の第6の実施の形態では、ノルム正規化部2122においてノルムが1になるような正規化を行ったが、ノルムが1以外の規定値となるように正規化を行ってもよい。さらに、ノルム正規化部2122を設けず、ノルム正規化を行わない構成であってもよい。この場合、クラスタリング部2123は、周波数正規化ベクトル $X'(f, \tau)$ のクラスタリ

ングを行うことになる。しかし、周波数正規化ベクトル $X'(f, \tau)$ は、ノルムが統一されていない。そこで、この場合のクラスタリング基準は、ベクトルがノルムを含めて似ているかどうかではなく、ベクトルの方向のみが似ているかどうかになる。これは、類似度を用いた評価になる。類似度の1つとしてコサイン距離

$$\cos \theta = |X'^H(f, \tau) \cdot \eta_k| / (\|X'(f, \tau)\| \cdot \|\eta_k\|)$$

を例示できる。ここで θ は、周波数正規化ベクトル $X'(f, \tau)$ と、セントロイド η_k のベクトルとがなす角度である。コサイン距離を用いる場合、クラスタリング部2123は、コサイン距離の総和

[0220] [数61]

$$U_i = \sum_{X_p'(f, \tau) \in C_i} |X_p'^H(f, \tau) \cdot \eta_i| / (\|X_p'(f, \tau)\| \cdot \|\eta_i\|)$$

を最小値化するクラスタを生成する。なお、セントロイド η_k は、各クラスタのメンバの平均として算出する。

また、上述した基準値 Q 、 Q' は互いに等しくてもよく、異なってもよい。

また、パラメータ d は、全てのセンサ q に対して同一であってもよいし、各センサ q に対応して複数の値を設定しておいてもよい。例えば、基準センサとセンサ q との距離を、それぞれのセンサ q に対応するパラメータ d の値としてもよい。

[0221] また、分離信号生成部2124において、

[0222] [数62]

$$Y_k(f, \tau) = \begin{cases} X_{Q'}(f, \tau) & X''(f, \tau) \in C_k \\ 0 & \text{otherwise} \end{cases}$$

とする代わりに、バイナリマスク

[0223] [数63]

$$M_k(f, \tau) = \begin{cases} 1 & X''(f, \tau) \in C_k \\ 0 & \text{otherwise} \end{cases}$$

を作成し、

$$Y_k(f, \tau) = M_k(f, \tau) X_{Q'}(f, \tau)$$

として分離信号ベクトル $Y(f, \tau)$ の k 番目の要素 $Y_k(f, \tau)$ を得ることとしてもよい。

さらに、上述の各実施の形態では、フーリエ変換及び逆フーリエ変換によって周波数領域と時間領域との間の変換を行うこととしたが、wavelet変換、DFTフィルタバンク、ポリフェイズフィルタバンクなどを用い、この変換を行うこととしてもよい(例えば、「R. E. Crochiere, L. R. Rabiner, "Multirate Digital Signal Processing." Eaglewood Cliffs, NJ: Printice-Hall, 1983. (ISBN 0-13-605162-6))。また、上述の各種の処理は、記載に従って時系列に実行されるのみならず、処理を実行する装置の処理能力あるいは必要に応じて並列的にあるいは個別に実行されてもよい。その他、本発明の趣旨を逸脱しない範囲で適宜変更が可能であることはいうまでもない。

[0224] また、上述の各実施の形態の構成をコンピュータによって実現する場合、各装置が有すべき機能の処理内容はプログラムによって記述される。そして、このプログラムをコンピュータで実行することにより、上記処理機能がコンピュータ上で実現される。

これらの処理内容を記述したプログラムは、コンピュータで読み取り可能な記録媒体に記録しておくことができる。コンピュータで読み取り可能な記録媒体としては、例えば、磁気記録装置、光ディスク、光磁気記録媒体、半導体メモリ等のようなものでもよいが、具体的には、例えば、磁気記録装置として、ハードディスク装置、フレキシブルディスク、磁気テープ等を、光ディスクとして、DVD (Digital Versatile Disc)、DVD-RAM (Random Access Memory)、CD-ROM (Compact Disc Read Only Memory)、CD-R (Recordable) / RW (ReWritable) 等を、光磁気記録媒体として、MO (Magnetooptical disc) 等を、半導体メモリとしてEEPROM (Electrically Erasable and Programmable-Read Only Memory) 等を用いることができる。

[0225] また、このプログラムの流通は、例えば、そのプログラムを記録したDVD、CD-ROM等の可搬型記録媒体を販売、譲渡、貸与等することによって行う。さらに、このプログラムをサーバコンピュータの記憶装置に格納しておき、ネットワークを介して、サーバコンピュータから他のコンピュータにそのプログラムを転送することにより、このプログラムを流通させる構成としてもよい。

また、上述した実施形態とは別の実行形態として、コンピュータが可搬型記録媒体から直接このプログラムを読み取り、そのプログラムに従った処理を実行することとし

てもよく、さらに、このコンピュータにサーバコンピュータからプログラムが転送されるたびに、逐次、受け取ったプログラムに従った処理を実行することとしてもよい。また、サーバコンピュータから、このコンピュータへのプログラムの転送は行わず、その実行指示と結果取得のみによって処理機能を実現する、いわゆるASP (Application Service Provider) 型のサービスによって、上述の処理を実行する構成としてもよい。なお、本形態におけるプログラムには、電子計算機による処理の用に供する情報であってプログラムに準ずるもの(コンピュータに対する直接の指令ではないがコンピュータの処理を規定する性質を有するデータ等)を含むものとする。

- [0226] また、この形態では、コンピュータ上で所定のプログラムを実行させることにより、本装置を構成することとしたが、これらの処理内容の少なくとも一部をハードウェア的に実現することとしてもよい。

産業上の利用可能性

- [0227] 本技術により、様々な妨害信号が発生する実環境において、目的の信号を精度良く取り出すことが可能となる。音信号に対する応用例としては、例えば、音声認識器のフロントエンドとして働く音声分離システムなどが挙げられる。話者とマイクが離れた位置にあり、マイクが話者の音声以外を集音してしまうような状況でも、そのようなシステムを使うことで、話者の音声のみを取り出して正しく音声を認識することができる。

請求の範囲

- [1] 複数の信号源から発せられた源信号の混合からなる混合信号を、当該源信号に分離する信号分離装置であって、
 複数のセンサで観測された前記混合信号を、周波数領域の混合信号に変換する周波数領域変換部と、
 前記周波数領域の混合信号を用いて生成された複素ベクトルを正規化し、当該複素ベクトルの周波数依存性を排除した正規化ベクトルを生成する正規化部と、
 前記正規化ベクトルをクラスタリングし、クラスタを生成するクラスタリング部と、を有する。
- [2] 請求項1に記載の信号分離装置であって、
 前記正規化部は、
 前記複素ベクトルが具備する特定の1つの要素を基準として当該複素ベクトルの各要素の偏角を正規化する第1正規化部と、
 前記第1正規化部で正規化された各要素の偏角を周波数に比例した値で除算する第2正規化部と、を有する。
- [3] 請求項2に記載の信号分離装置であって、
 前記正規化部は、
 前記第2正規化部で正規化された各要素からなるベクトルのノルムを所定の値に正規化する第3正規化部をさらに有する。
- [4] 複数の信号源から発せられた源信号の混合からなる混合信号を、当該源信号に分離する信号分離装置であって、
 複数のセンサで観測された前記混合信号を、周波数領域の混合信号に変換する周波数領域変換部と、
 前記周波数領域の混合信号を用いて、周波数毎に分離行列を算出する分離行列算出部と、
 前記分離行列の一般化逆行列を算出する逆行列算出部と、
 前記一般化逆行列を構成する基底ベクトルの正規化を行い、正規化基底ベクトルを算出する基底ベクトル正規化部と、

前記正規化基底ベクトルをクラスタリングし、クラスタを生成するクラスタリング部と、
前記クラスタの中心ベクトルと前記正規化基底ベクトルとを用い、前記分離行列の
要素を並び替えるための順列を算出する順列算出部と、を有する。

- [5] 請求項4記載の信号分離装置であって、
前記基底ベクトル正規化部は、
前記基底ベクトルの周波数依存性を排除する正規化を行う。
- [6] 請求項5記載の信号分離装置であって、
前記基底ベクトルの周波数依存性を排除する正規化は、
前記基底ベクトルが具備する特定の1つの要素を基準として当該基底ベクトルの各
要素の偏角を正規化し、さらに、各要素の偏角を周波数に比例した値で除算する正
規化である。
- [7] 請求項5記載の信号分離装置であって、
前記基底ベクトルの周波数依存性を排除する正規化は、
前記基底ベクトル $A_p(f)$ ($p=1, \dots, N$, N は前記信号源の数)の各要素 $A_{qp}(f)$ ($q=1, \dots, M$, M は前記混合信号を観測するセンサの数)に対し、 \exp をネピア数とし、 $\arg[\cdot]$ を偏角とし、 f を周波数とし、 j を虚数単位とし、 c を信号の伝達速度とし、 Q を M 以下の自然数から選択された基準値とし、 d を実数とした場合における、

[数64]

$$A_{qp}'(f) = |A_{qp}(f)| \exp \left[j \frac{\arg[A_{qp}(f)/A_{Qp}(f)]}{4fc^{-1}d} \right]$$

の演算によって行われる正規化である。

- [8] 請求項7記載の信号分離装置であって、
前記 d は、
前記要素 $A_{Qp}(f)$ に対応する基準センサと他のセンサとの最大距離 d_{\max} である。
- [9] 請求項4記載の信号分離装置であって、
前記基底ベクトル正規化部は、
前記基底ベクトルの周波数依存性を排除する正規化と、ノルムを規定値にする正
規化とを行う。

- [10] 複数の信号源から発せられた源信号の混合からなる混合信号を、当該源信号に分離する信号分離装置であって、
複数のセンサで観測された前記混合信号を周波数領域の混合信号に変換する周波数領域変換部と、
前記周波数領域の混合信号を用いて、周波数毎に分離行列を算出する分離行列算出部と、
前記分離行列の一般化逆行列を算出する逆行列算出部と、
前記一般化逆行列を構成する基底ベクトルの正規化を行い、正規化基底ベクトルを算出する基底ベクトル正規化部と、
前記正規化基底ベクトルをクラスタリングし、クラスタを生成するクラスタリング部と、
前記周波数領域の混合信号と前記分離行列或いはその行を並び替えた分離行列とから求めた分離信号のエンベロープと、前記クラスタの中心ベクトルと、前記正規化基底ベクトルとを用い、前記分離行列の要素を並び替えるための順列を算出する順列算出部と、を有する。
- [11] 複数の信号源から発せられた源信号の混合からなる混合信号を、当該源信号に分離する信号分離装置であって、
複数のセンサで観測された前記混合信号を周波数領域の混合信号に変換する周波数領域変換部と、
前記周波数領域の混合信号を用い、周波数毎に分離行列と分離信号とを算出する信号分離部と、
前記分離行列の一般化逆行列の列である基底ベクトルを正規化し、その正規化された基底ベクトルをクラスタリングし、そのクラスタの分散を指標として前記分離信号から、目的信号を含む選択信号を選択する目的信号選択部と、を有する。
- [12] 請求項11に記載の信号分離装置であって、
更に、前記周波数領域の混合信号と前記基底ベクトルとを用いて時間周波数マスクを生成するマスク生成部と、
前記時間周波数マスクを前記目的信号選択部で選択された前記選択信号に適用し、マスク処理選択信号を生成するマスキング部と、を有する。

- [13] 請求項12に記載の信号分離装置であって、
 前記マスク生成部は、
 前記周波数領域の混合信号を用いて白色化行列を生成する白色化行列生成部と、
 前記白色化行列を用い、前記周波数領域の混合信号を要素とする混合信号ベクトルを変換した白色化混合信号ベクトルと、前記基底ベクトルを変換した白色化基底ベクトルとを算出する白色化部と、
 前記白色化混合信号ベクトルと前記白色化基底ベクトルとがなす角度を時間周波数毎に算出する角度計算部と、
 前記角度を要素とする関数である前記時間周波数マスクを生成する関数演算部とを有する。
- [14] 請求項13に記載の信号分離装置であって、
 前記白色化行列は、
 周波数を f とし、離散時間を τ とし、前記混合信号ベクトルを $X(f, \tau)$ とし、ベクトル $*$ を時間平均したベクトルを $\langle * \rangle$ とし、ベクトル $*$ の複素共役転置ベクトルを $*^H$ とし、 $R(f) = \langle X(f, \tau) \cdot X(f, \tau)^H \rangle$ とした場合における $V(f) = R(f)^{-1/2}$ であり、
 前記白色化部は、
 $Z(f, \tau) = V(f) \cdot X(f, \tau)$ の演算により、前記白色化混合信号ベクトル $Z(f, \tau)$ を算出し、前記基底ベクトルを $A(f)$ とした場合における $B(f) = V(f) \cdot A(f)$ の演算により、前記白色化基底ベクトル $B(f)$ を算出し、
 前記角度計算部は、
 ベクトル $*$ の絶対値を $|*|$ とし、ベクトル $*$ のノルムを $\|*\|$ とした場合における $\theta(f, \tau) = \cos^{-1}(|B^H(f) \cdot Z(f, \tau)| / \|B(f)\| \cdot \|Z(f, \tau)\|)$ の演算により、前記角度 $\theta(f, \tau)$ を算出し、
 前記関数演算部は、
 α, g, θ を実数とした場合におけるロジスティック関数 $M(\theta(f, \tau)) = \alpha / (1 + e^{g \cdot (\theta(f, \tau) - \theta_T)})$ を前記時間周波数マスクとして算出する。
- [15] 請求項12に記載の信号分離装置であって、

前記マスク生成部は、

前記周波数領域の混合信号を用いて生成された混合信号ベクトル $X(f, \tau)$ を、周波数に依存しない周波数正規化ベクトル $X'(f, \tau)$ に正規化する周波数正規化部と、

前記周波数正規化ベクトル $X'(f, \tau)$ を、ノルムが所定の値になるノルム正規化ベクトル $X''(f, \tau)$ に正規化する第1ノルム正規化部と、

前記選択信号に対応するセントロイド η を抽出するセントロイド選択部と、

前記選択信号に対応するセントロイド η を、そのノルムが前記所定の値になるノルム正規化セントロイド η' に正規化する第2ノルム正規化部と、

前記周波数正規化ベクトル $X'(f, \tau)$ と前記ノルム正規化セントロイド η' との距離の自乗 $DS(f, \tau)$ を算出する自乗距離算出部と、

前記距離の自乗 $DS(f, \tau)$ を要素とする関数を用いた前記時間周波数マスクを生成する

関数生成部と、を有する。

[16] 請求項11に記載の信号分離装置であって、

前記目的信号選択部は、

前記基底ベクトルの周波数依存性を排除する正規化を行う。

[17] 請求項16記載の信号分離装置であって、

前記基底ベクトルの周波数依存性を排除する正規化は、

前記基底ベクトルが具備する特定の1つの要素を基準として当該基底ベクトルの各要素の偏角を正規化し、さらに、各要素の偏角を周波数に比例した値で除算する正規化である。

[18] 請求項17記載の信号分離装置であって、

前記基底ベクトルの周波数依存性を排除する正規化は、

前記基底ベクトル $A_p(f)$ (p は自然数)の各要素 $A_{pq}(f)$ ($q=1, \dots, M$, M は前記混合信号を観測するセンサの数)に対し、 \exp をネピア数とし、 $\arg[\cdot]$ を偏角とし、 f を周波数とし、 j を虚数単位とし、 c を信号の伝達速度とし、 Q を M 以下の自然数から選択された基準値とし、 d を実数とした場合における、

[数65]

$$A_{qp}'(f) = |A_{qp}(f)| \exp \left[j \frac{\arg[A_{qp}(f)/A_{Qp}(f)]}{4fc^{-1}d} \right]$$

の演算によって行われる正規化である。

- [19] 請求項18記載の信号分離装置であって、
前記実数dは、
前記基準値Qに対応する基準センサと他のセンサとの最大距離 d_{\max} である。
- [20] 請求項11記載の信号分離装置であって、
前記目的信号選択部は、
前記基底ベクトルの周波数依存性を排除する正規化と、ノルムを規定値にする正規化とを行う。
- [21] 請求項11に記載の信号分離装置であって、
前記目的信号選択部は、
分散が最小となるクラスタを選択し、選択されたクラスタに対応する前記分離信号を前記選択信号として選択する。
- [22] 複数の信号源から発せられた源信号の混合からなる混合信号を、当該源信号に分離する信号分離装置であって、
複数のセンサで観測された前記混合信号を周波数領域の混合信号に変換する周波数領域変換部と、
前記周波数領域の混合信号から構成される混合信号ベクトルの正規化を行い、正規化ベクトルを算出するベクトル正規化部と、
前記正規化ベクトルをクラスタリングし、クラスタを生成するクラスタリング部と、
k番目の前記クラスタに属する前記正規化ベクトルの時間周波数に対応する前記混合信号ベクトルから所定番目の要素を抽出し、これをk番目の要素とした分離信号ベクトルを生成する分離信号生成部と、を有する。
- [23] 請求項22に記載の信号分離装置であって、
前記ベクトル正規化部は、
前記周波数領域の混合信号から構成される混合信号ベクトルの周波数依存性を排除する正規化を行う。

- [24] 請求項23記載の信号分離装置であって、
 前記混合信号ベクトルの周波数依存性を排除する正規化は、
 前記混合信号ベクトルが具備する特定の1つの要素を基準として当該基底ベクトルの各要素の偏角を正規化し、さらに、各要素の偏角を周波数に比例した値で除算する正規化である。
- [25] 請求項24記載の信号分離装置であって、
 前記混合信号ベクトルの周波数依存性を排除する正規化は、
 M を前記センサの数とし、 $q=1, \dots, M$ とし、前記混合信号ベクトルの各要素を $X_q(f, \tau)$ とし、 \exp をネピア数とし、 $\arg[\cdot]$ を偏角とし、 j を虚数単位とし、 c を信号の伝達速度とし、 Q を M 以下の自然数から選択された値とし、 d を実数とし、 f を周波数とし、 τ を離散時間とした場合における、
 [数66]

$$X_q'(f, \tau) = |X_q(f, \tau)| \exp \left[j \frac{\arg[X_q(f, \tau)/X_Q(f, \tau)]}{4fc^{-1}d} \right]$$

の演算によって行われる。

- [26] 請求項25記載の信号分離装置であって、
 前記 d は、
 前記要素 $X_q(f, \tau)$ に対応するセンサと他のセンサとの最大距離 d_{\max} である。
- [27] 請求項22に記載の信号分離装置であって、
 前記ベクトル正規化部は、
 前記混合信号ベクトルの周波数依存性を排除する正規化と、ノルムを規定値にする正規化とを行う。
- [28] 複数の信号源から発せられた源信号の混合からなる混合信号を、当該源信号に分離する信号分離方法であって、
 複数のセンサで観測された前記混合信号を周波数領域の混合信号に変換して出力する過程と、
 入力された前記周波数領域の混合信号を用い、周波数毎に分離行列を算出して出力する過程と、

入力された前記分離行列の一般化逆行列を算出して出力する過程と、

入力された前記一般化逆行列を構成する基底ベクトルの正規化を行い、正規化基底ベクトルを算出して出力する過程と、

入力された前記正規化基底ベクトルをクラスタリングし、クラスタを生成して出力する過程と、

入力された前記クラスタの中心ベクトルと前記正規化基底ベクトルとを用い、前記分離行列の要素を並び替えるための順列を算出して出力する過程と、を有する。

- [29] 複数の信号源から発せられた源信号の混合からなる混合信号を、当該源信号に分離する信号分離方法であって、

複数のセンサで観測された前記混合信号を周波数領域の混合信号に変換して出力する過程と、

入力された前記周波数領域の混合信号を用い、周波数毎に分離行列と分離信号とを算出して出力する過程と、

入力された前記分離行列の一般化逆行列の列である基底ベクトルを正規化し、その正規化された基底ベクトルをクラスタリングし、そのクラスタの分散を指標として、入力された前記分離信号から、目的信号を含む選択信号を選択する過程と、を有する。

。

- [30] 複数の信号源から発せられた源信号の混合からなる混合信号を、当該源信号に分離する信号分離方法であって、

複数のセンサで観測された前記混合信号を周波数領域の混合信号に変換して出力する過程と、

入力された前記周波数領域の混合信号から構成される混合信号ベクトルの正規化を行い、正規化ベクトルを算出して出力する過程と、

入力された前記正規化ベクトルをクラスタリングし、クラスタを生成して出力する過程と、

k番目の前記クラスタに属する前記正規化ベクトルの時間周波数に対応する前記混合信号ベクトルから所定番目の要素を抽出し、これをk番目の要素とした分離信号ベクトルを生成して出力する過程と、を有する。

- [31] 請求項1から27の何れかに記載の信号分離装置としてコンピュータを機能させるための信号分離プログラム。
- [32] 請求項31記載の信号分離プログラムを格納したコンピュータ読み取り可能な記録媒体。

[図1]

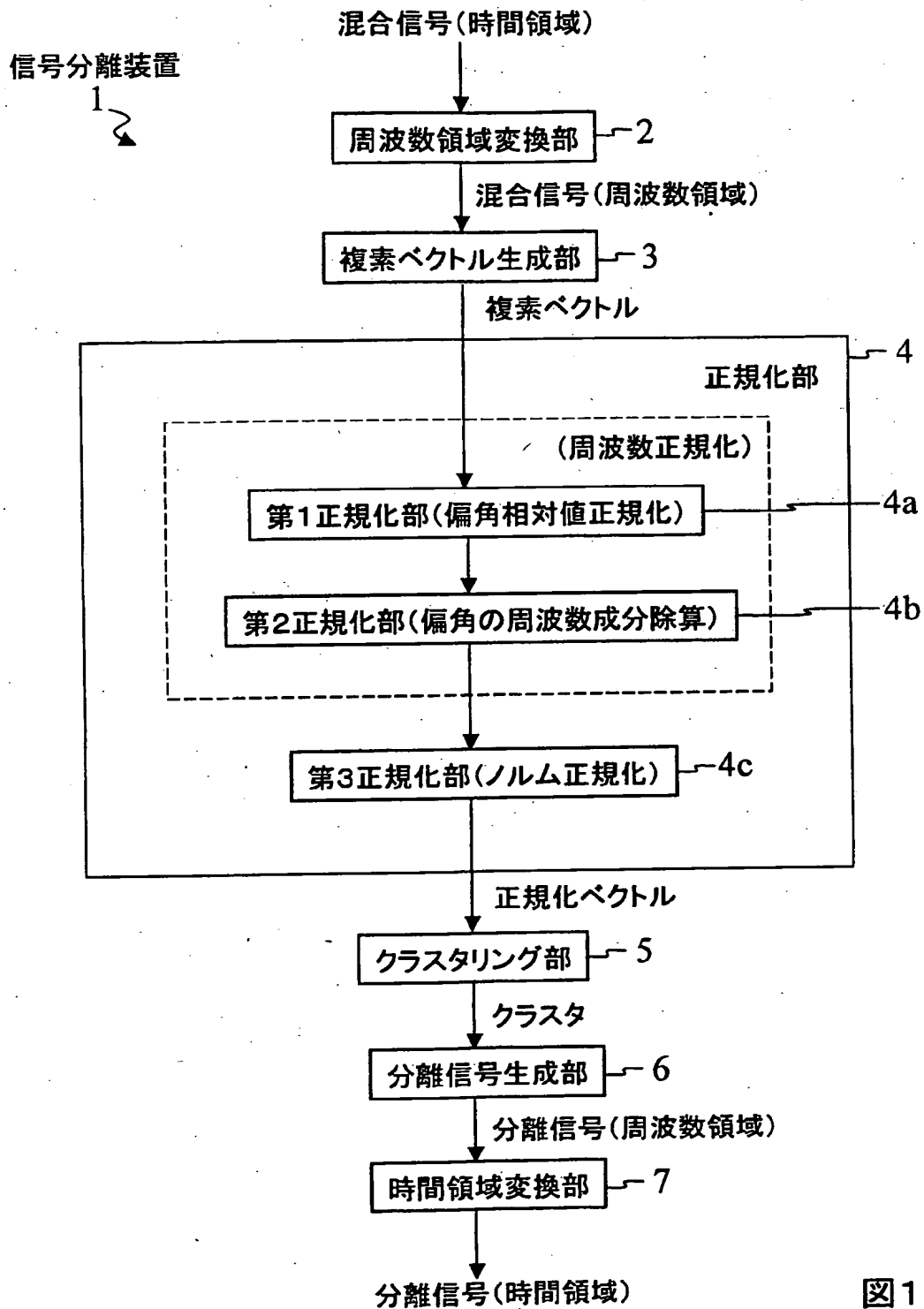


図1

[図2]

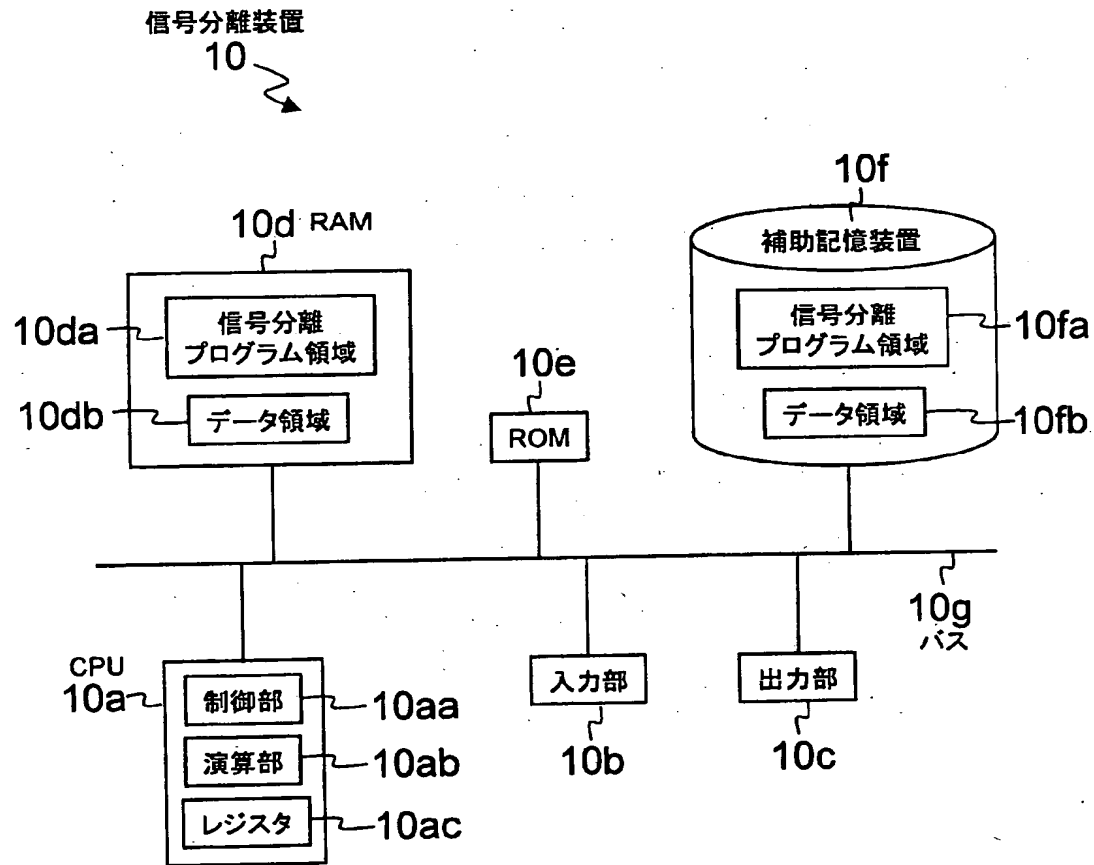


図2

[図3]

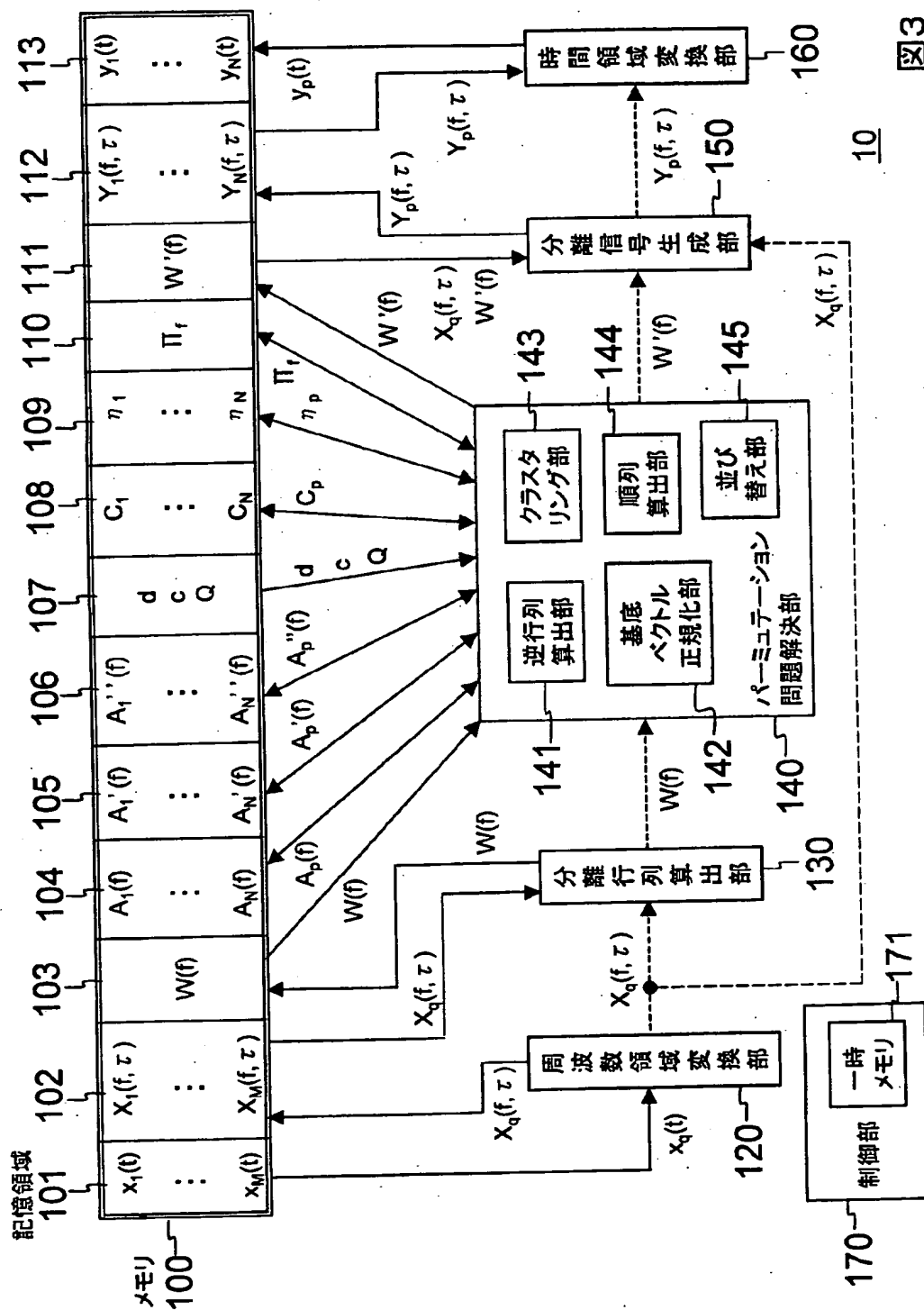


図3

[図4]

図4A

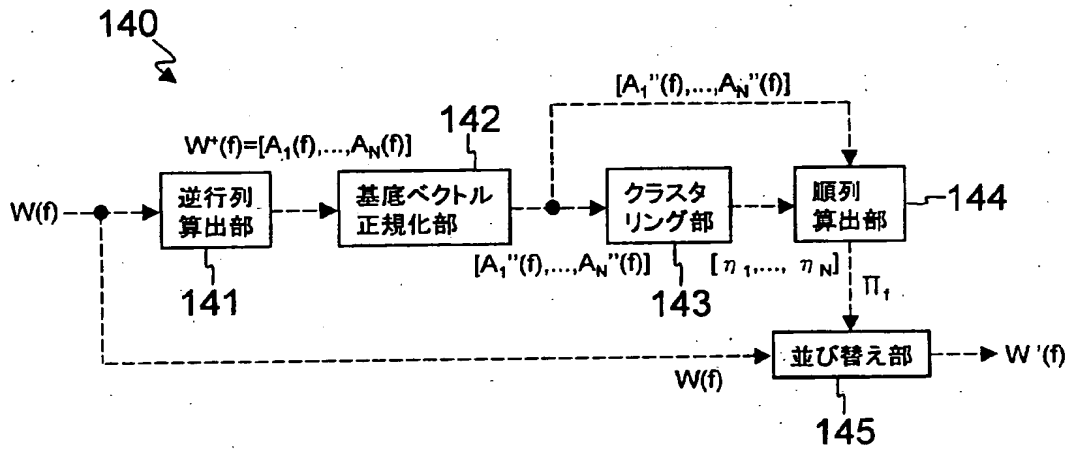
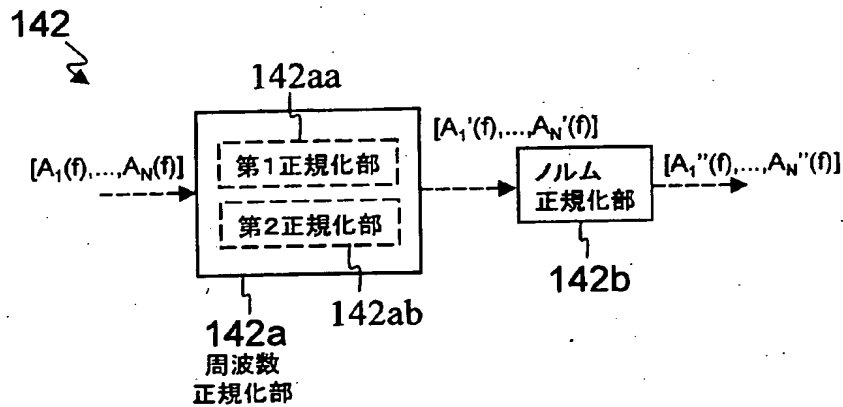


図4B



[図5]

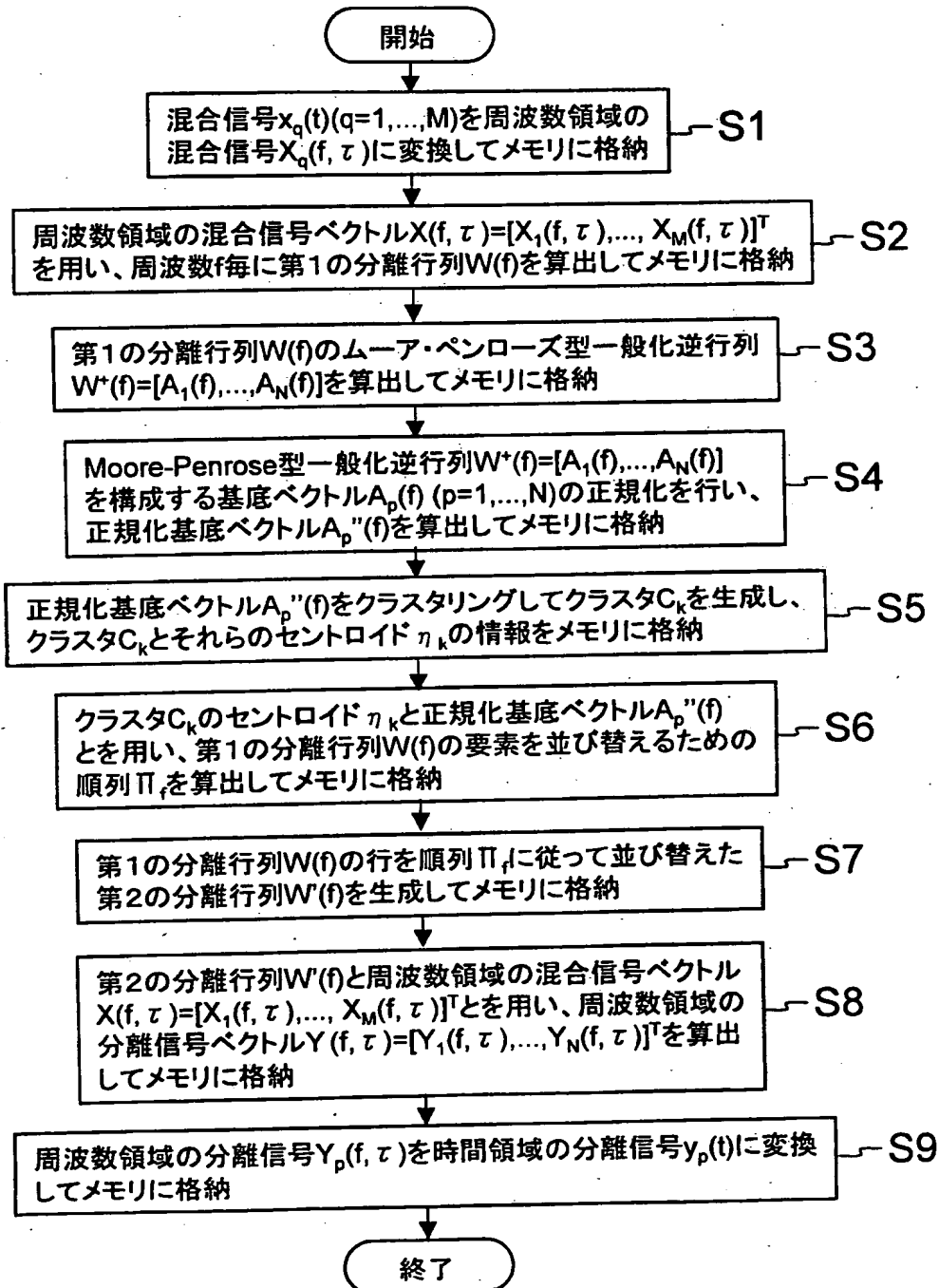


図5

[図6]

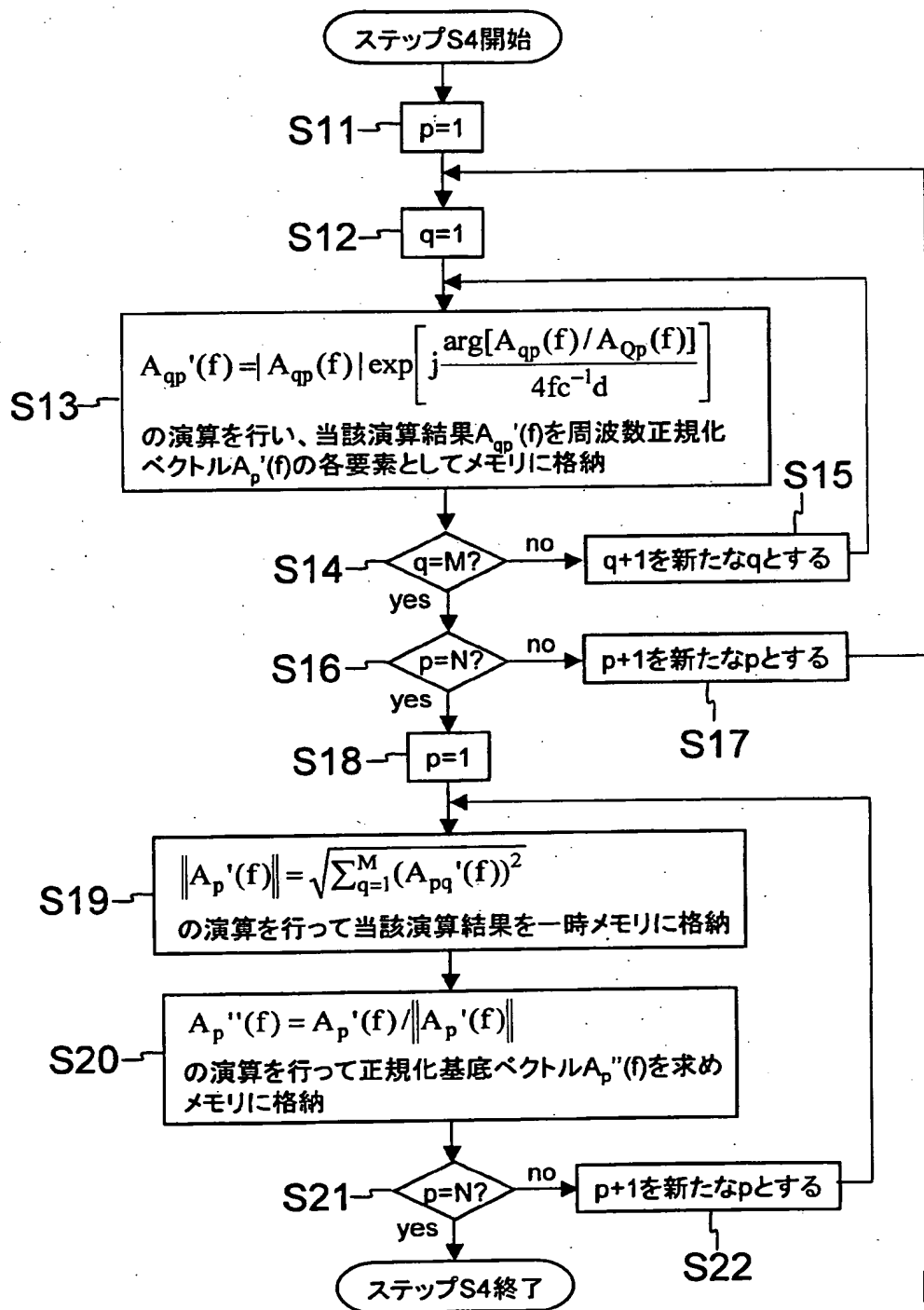


図6

[図7]

図7A

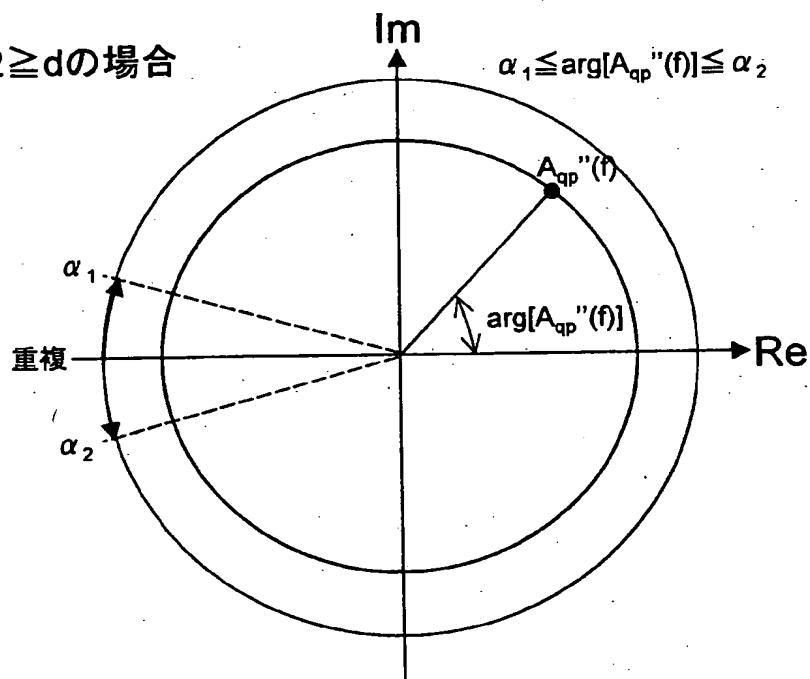
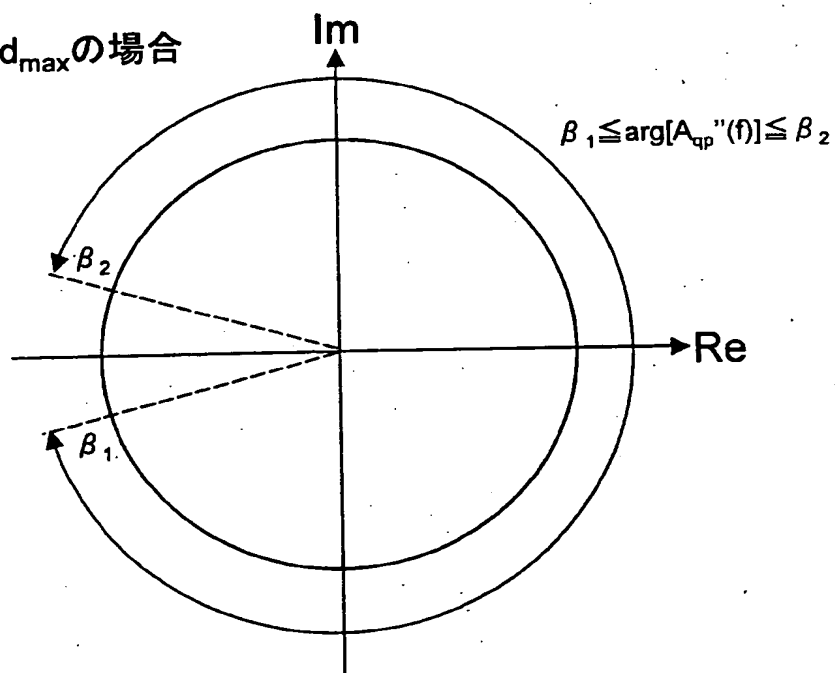
 $d_{\max}/2 \geq d$ の場合

図7B

 $d_{\max}/2 < d < d_{\max}$ の場合

[図8]

図8A

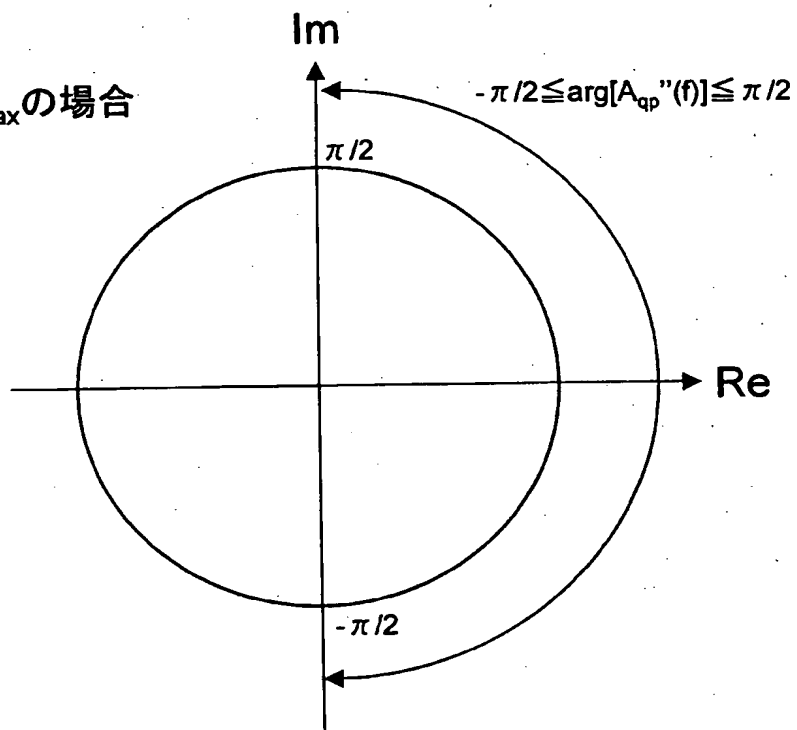
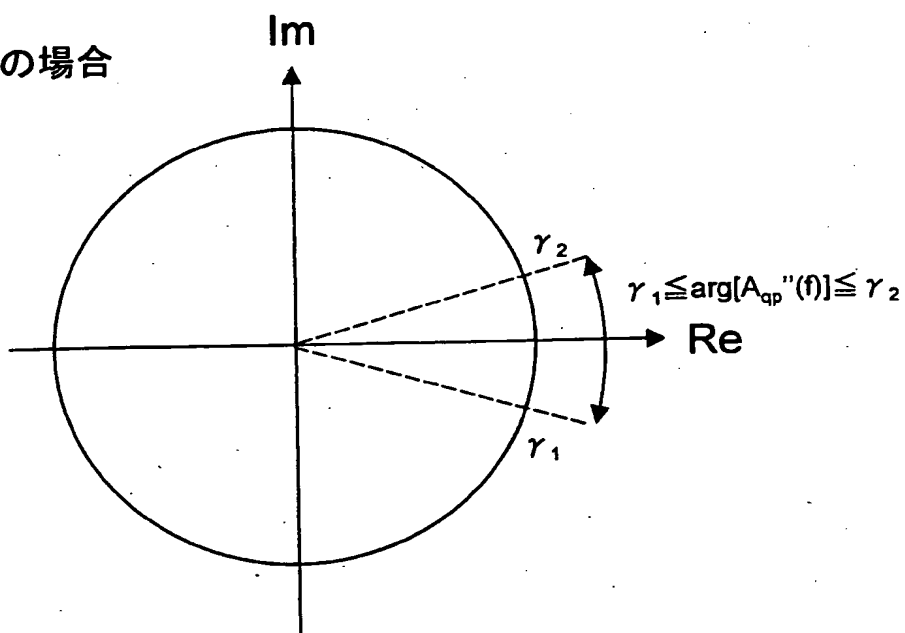
 $d=d_{\max}$ の場合

図8B

 $d > d_{\max}$ の場合

[図9]

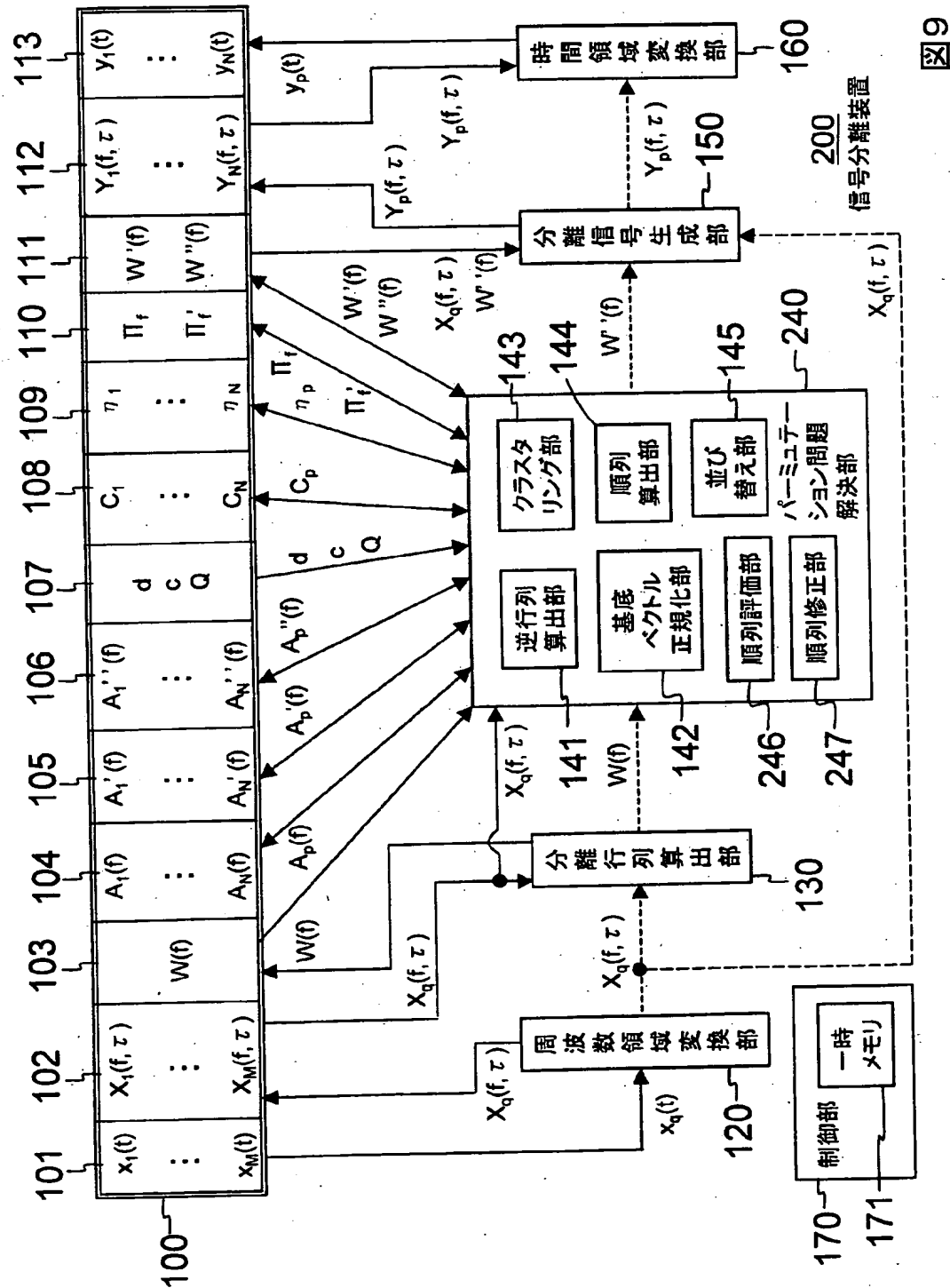


図9

[図10]

図10A

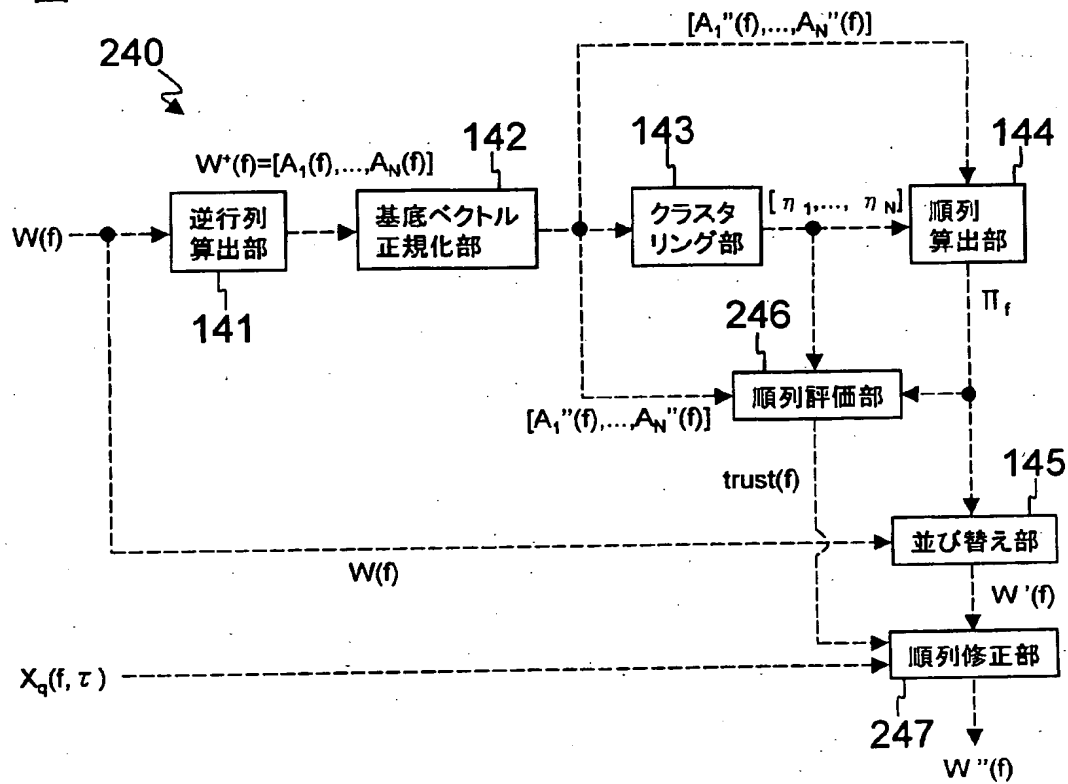
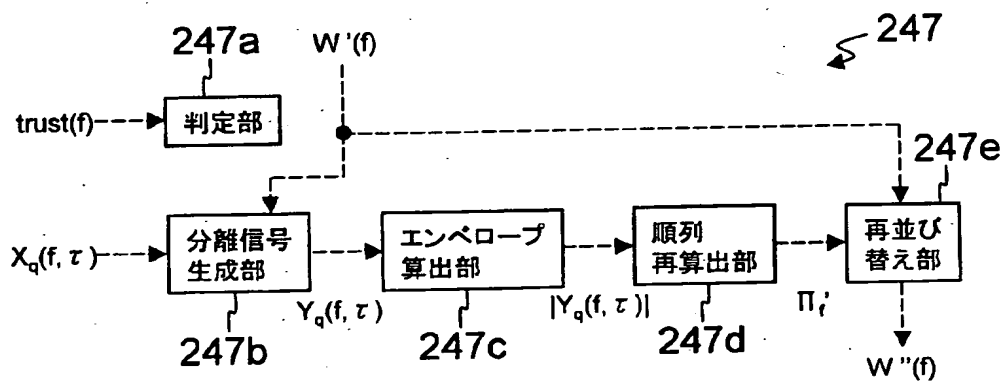


図10B



[図11]

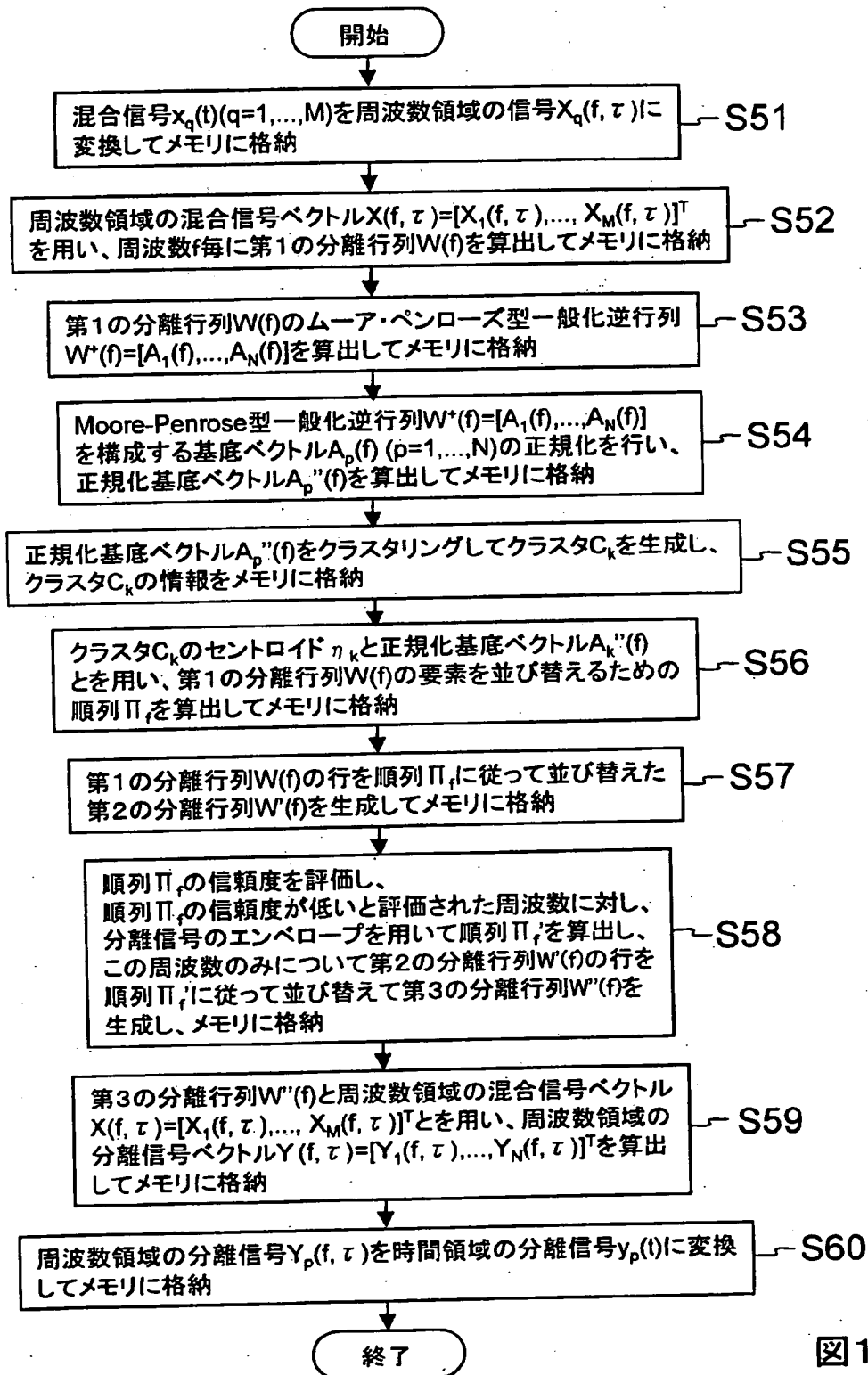


図11

[図12]

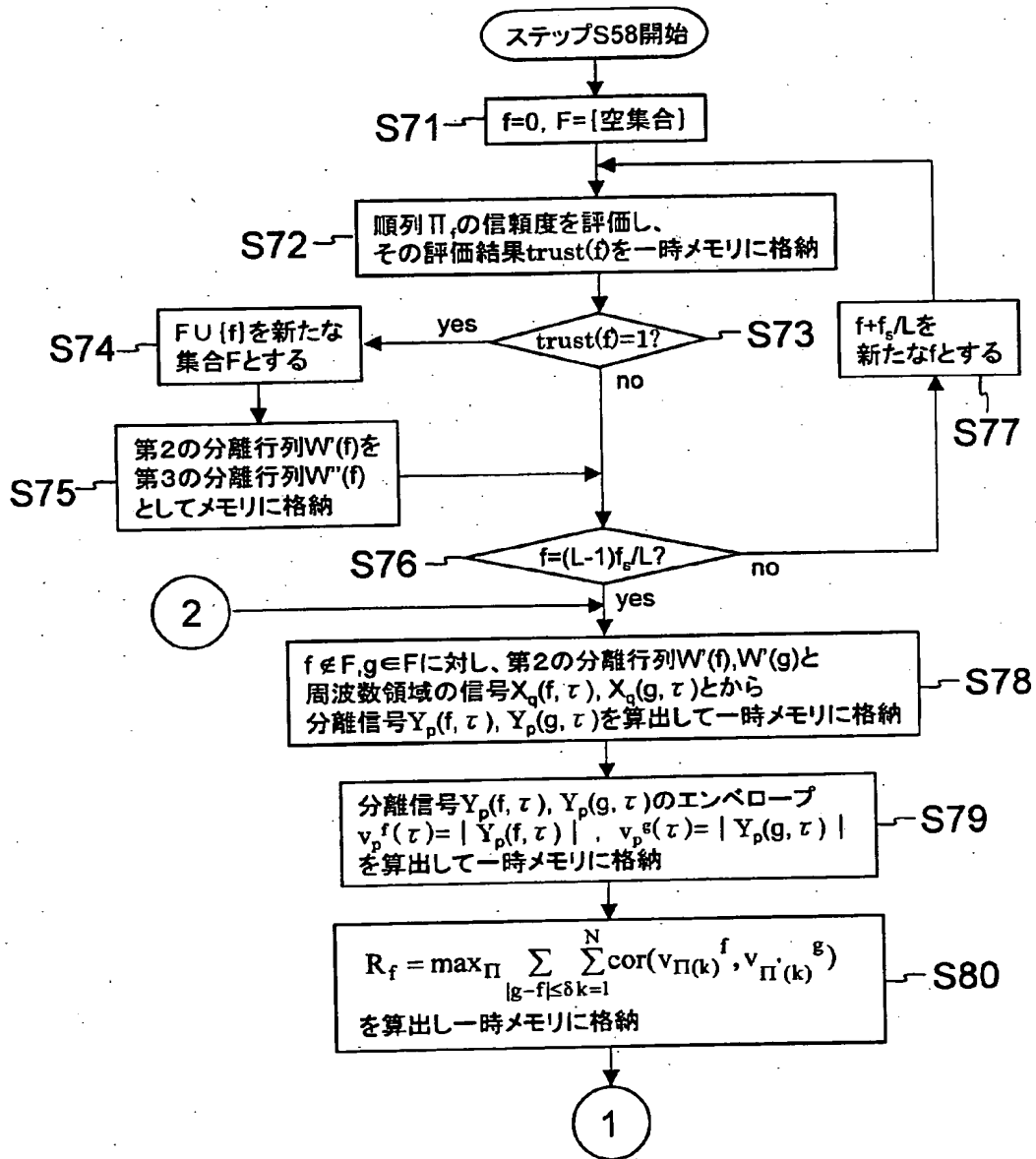


図12

[図13]

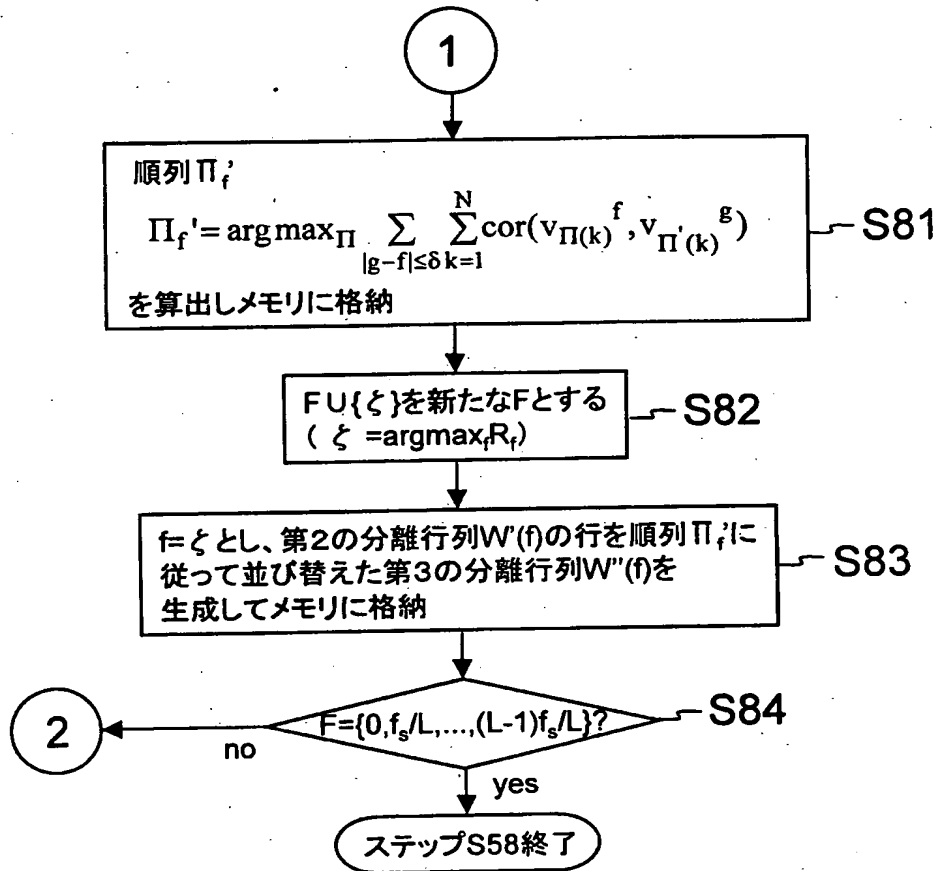


図13

[図14]

図14A

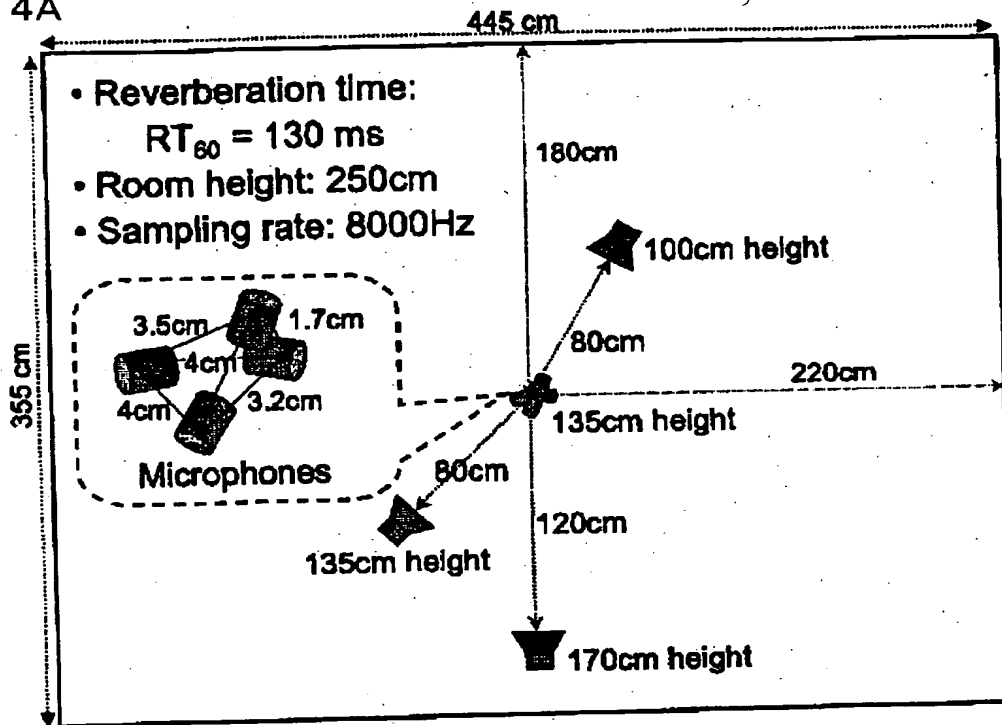
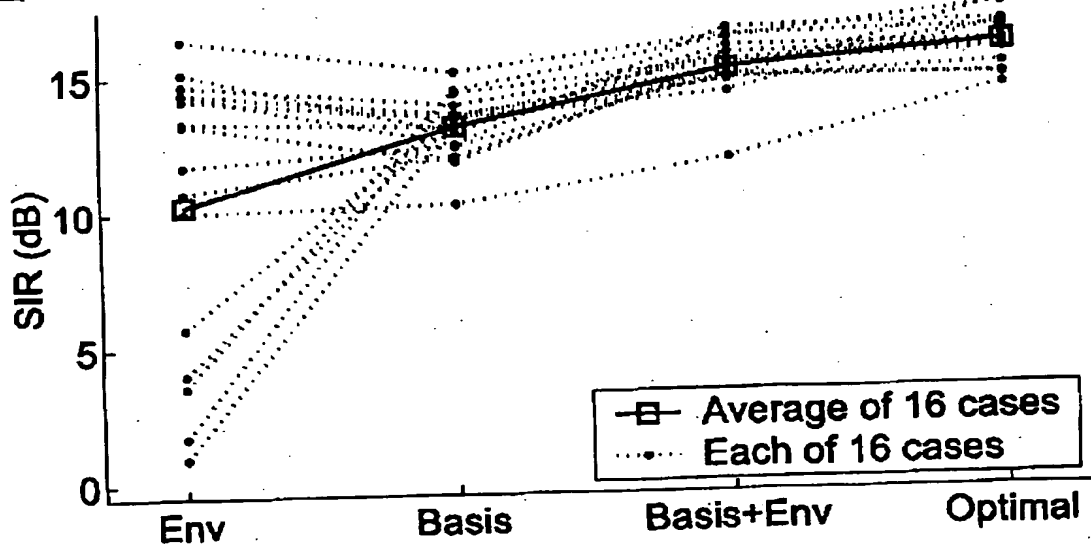


図14B



[図15]

図15A

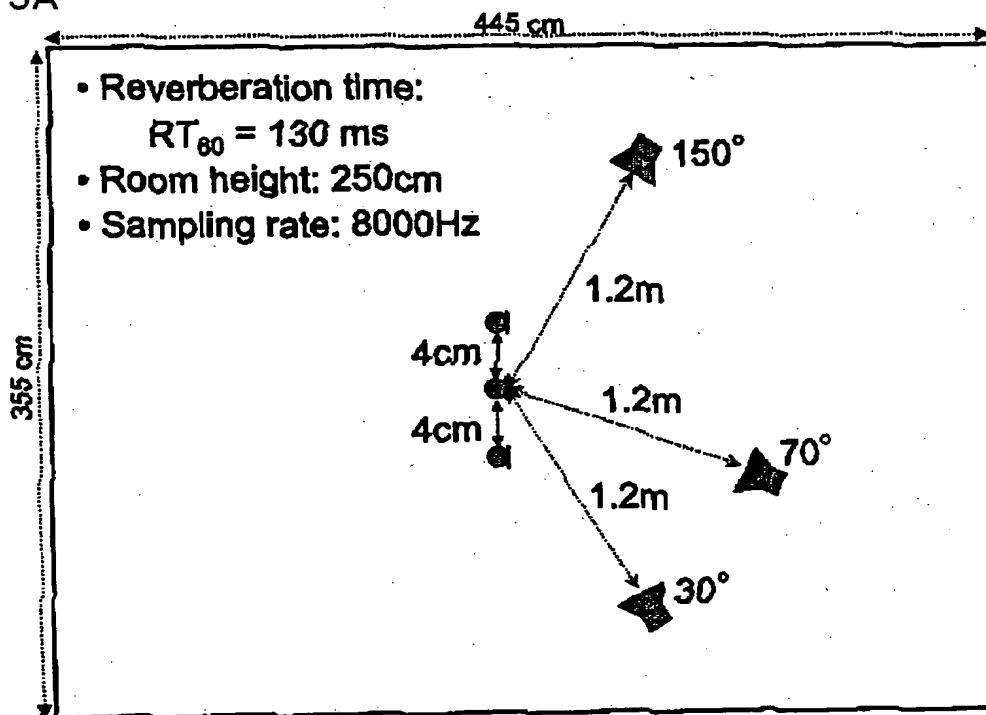
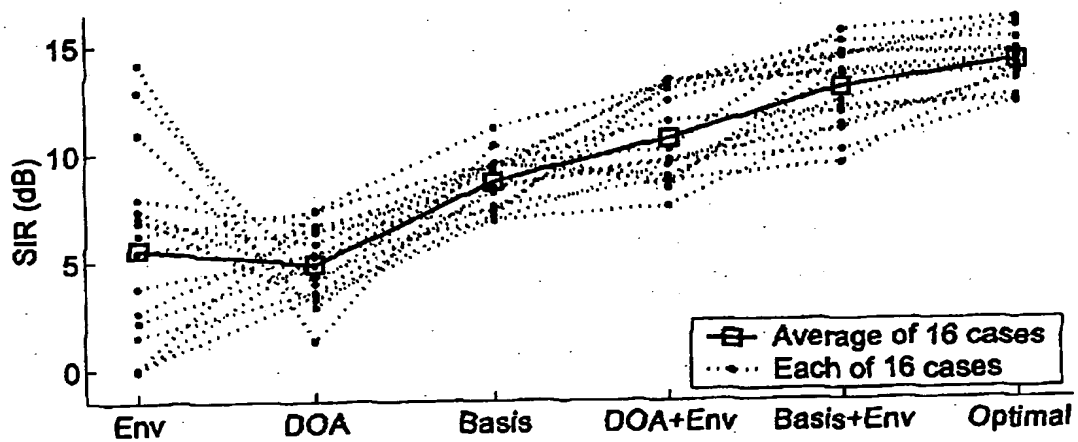
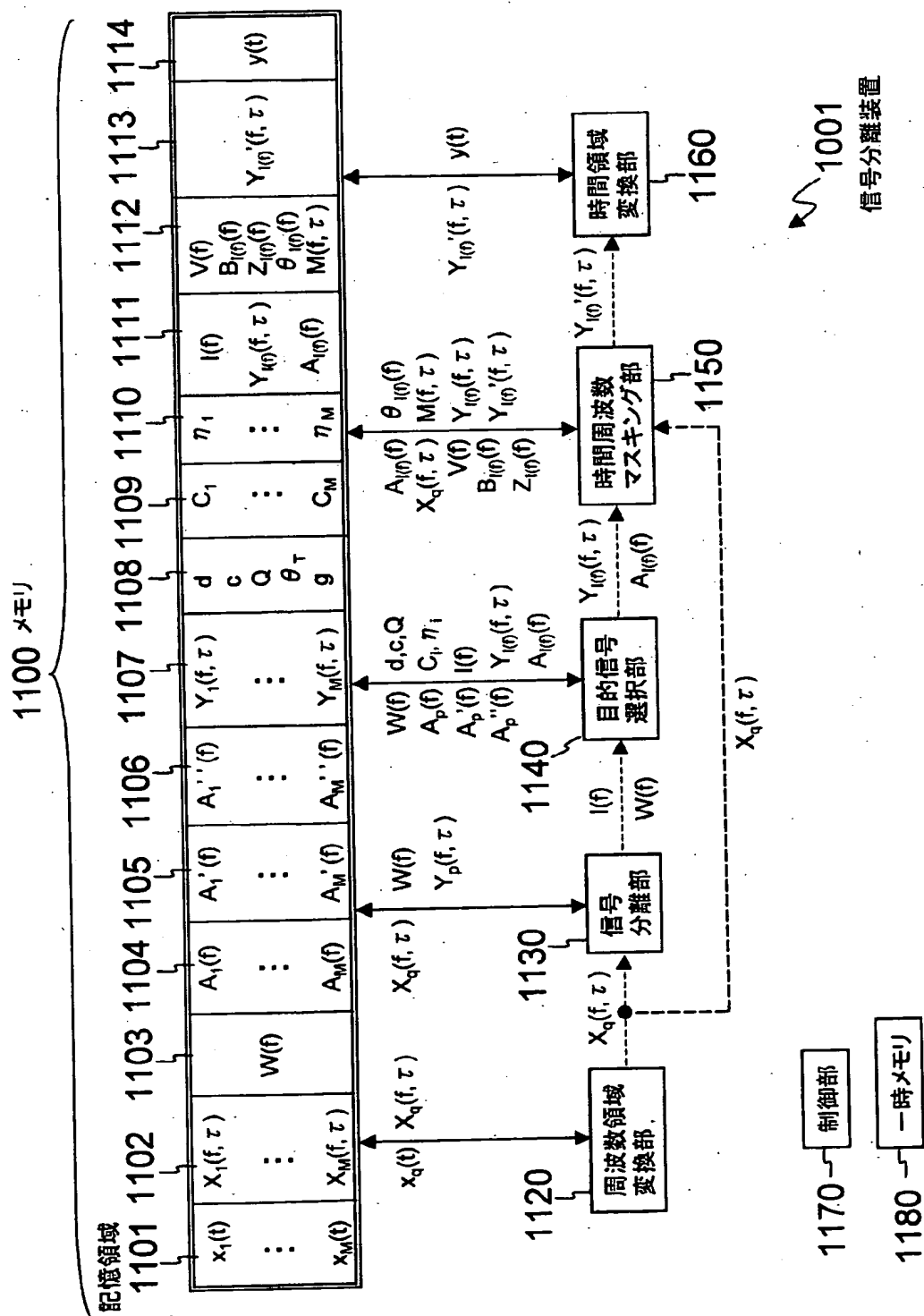


図15B



[図16]



[図17]

图 17A

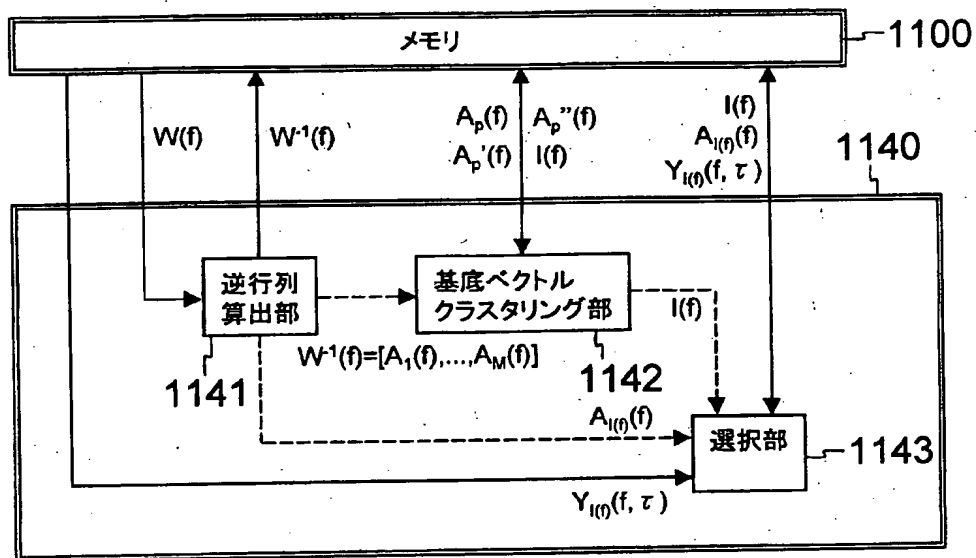
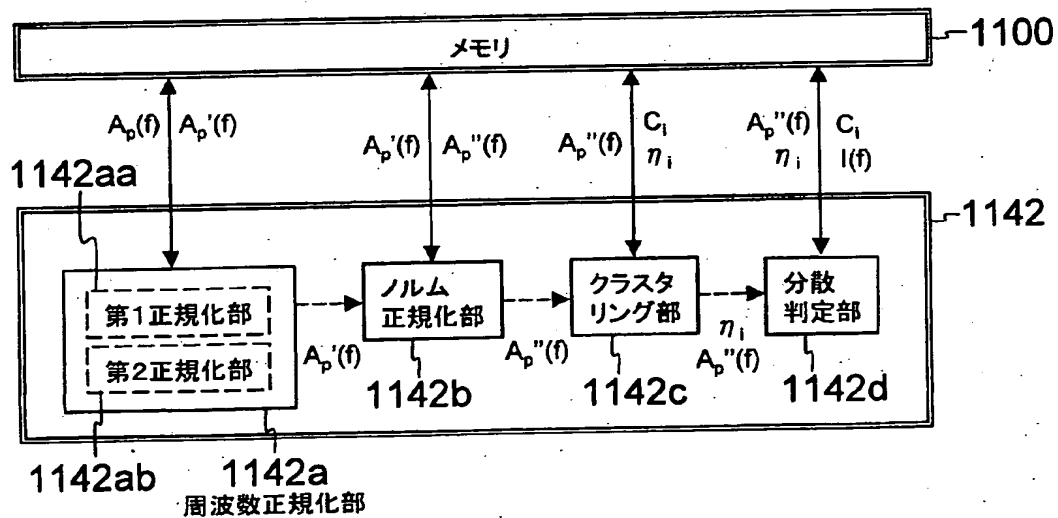


图17B



[図18]

図18A

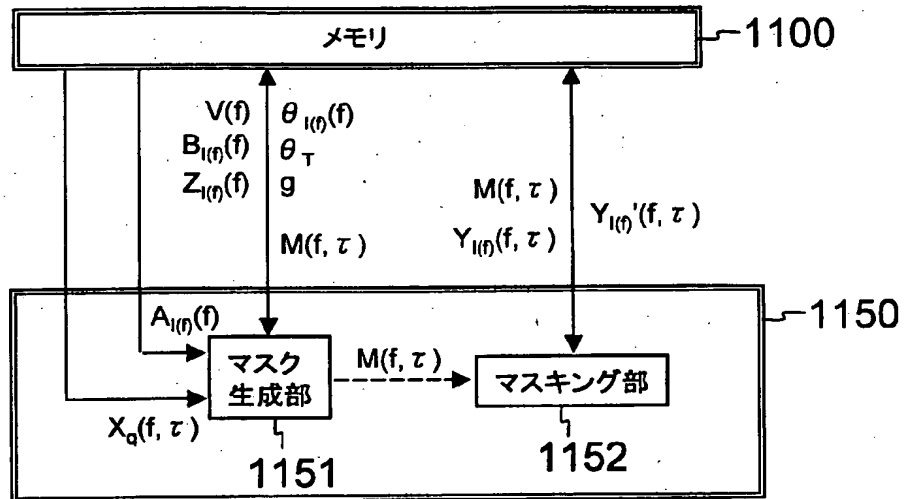
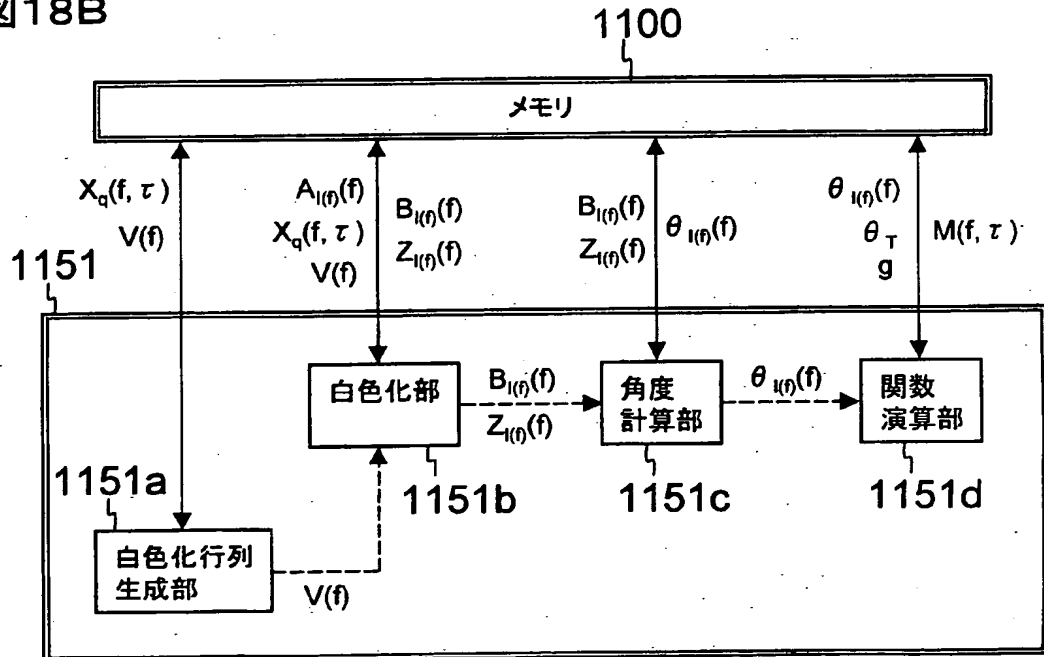


図18B



[図19]

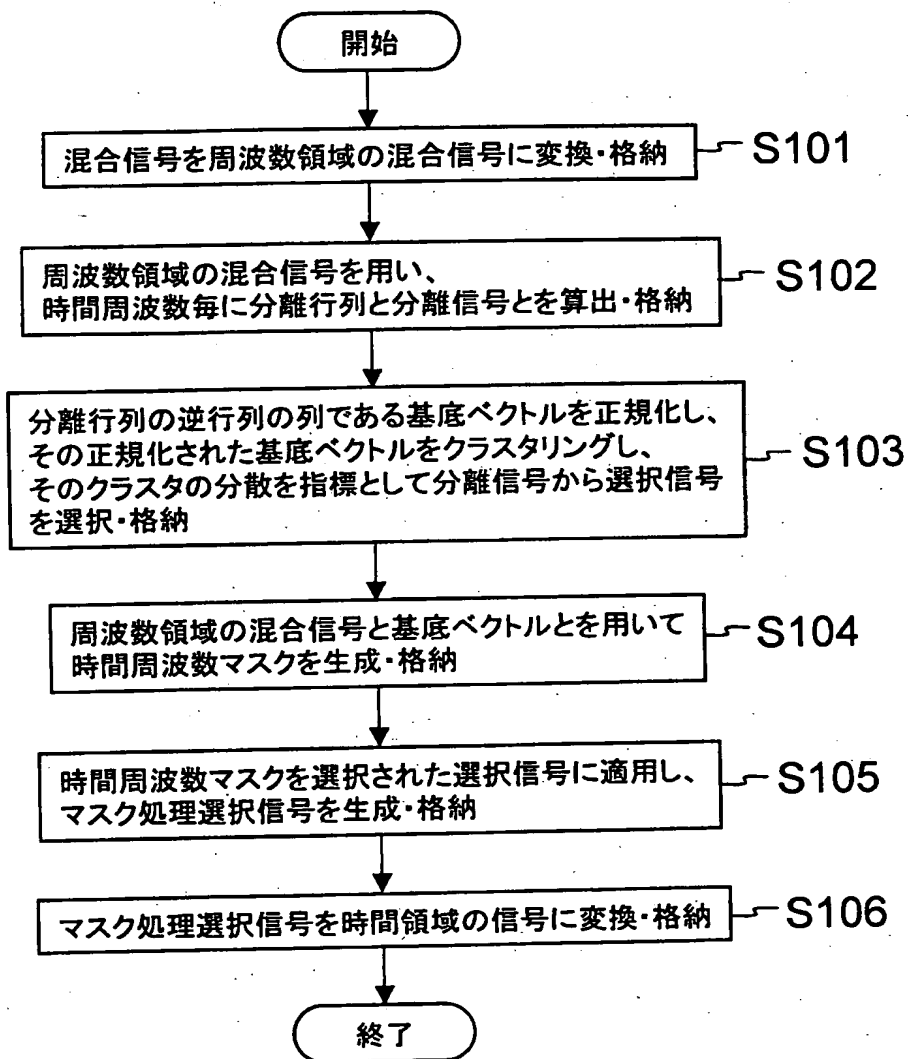


図19

[図20]

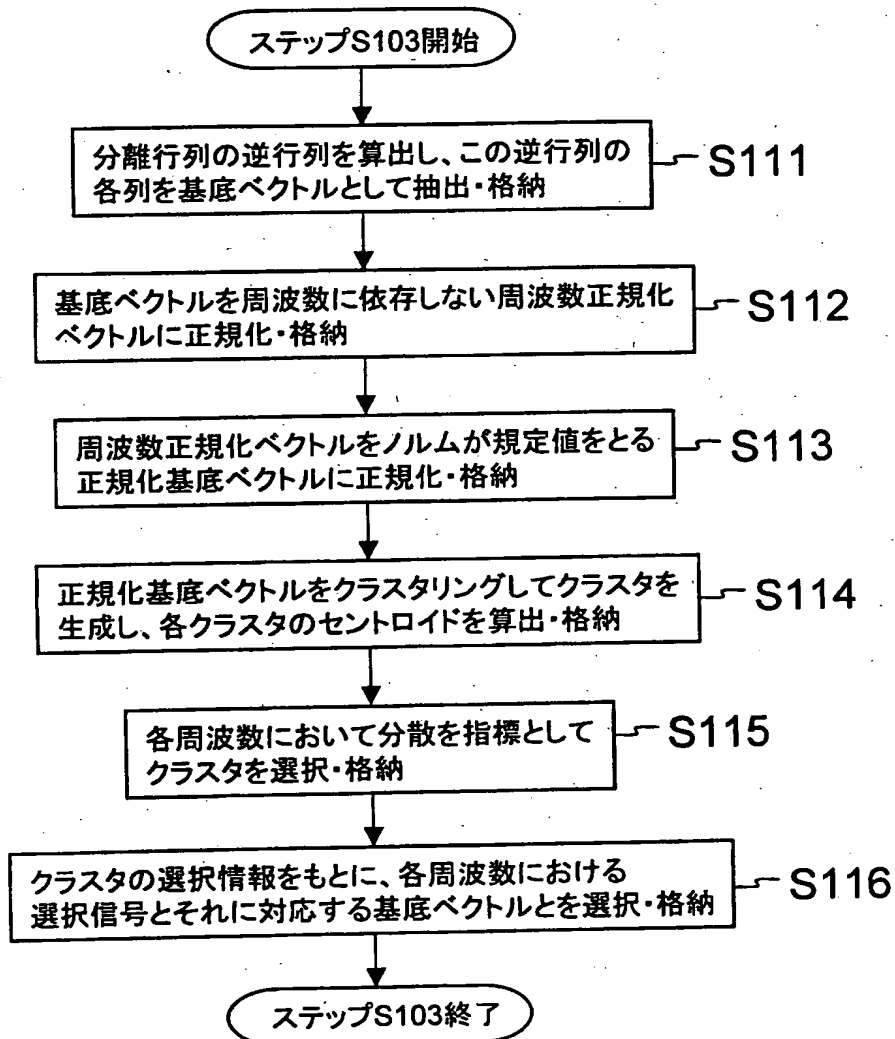


図20

[図21]

図21A

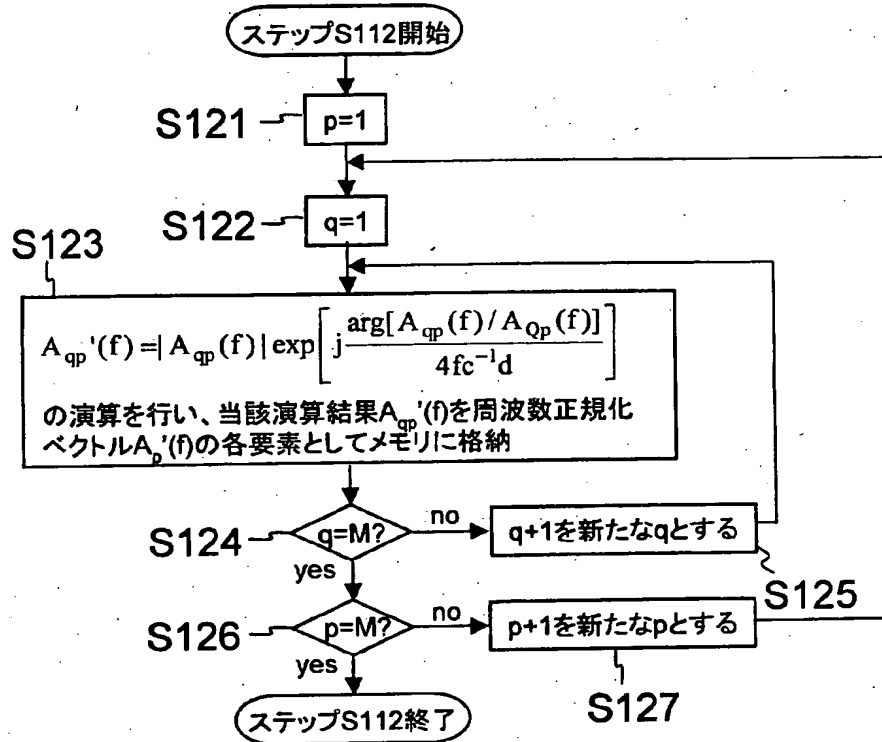
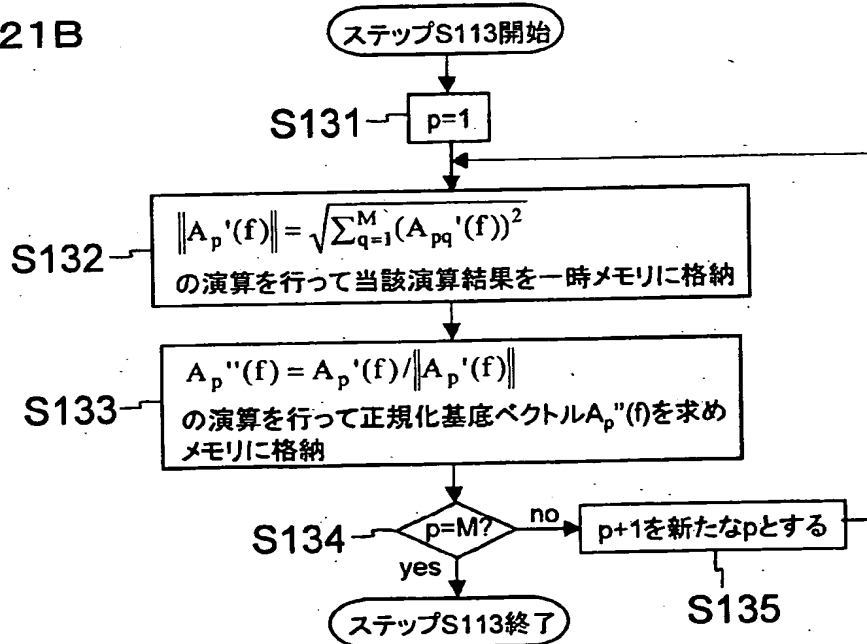


図21B



[図22]

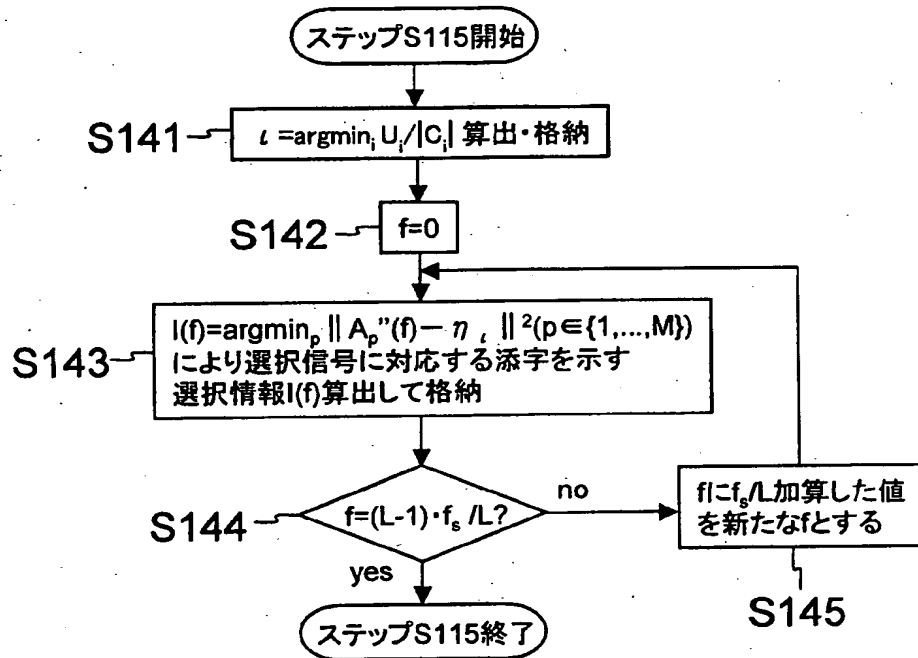


図22

[図23]

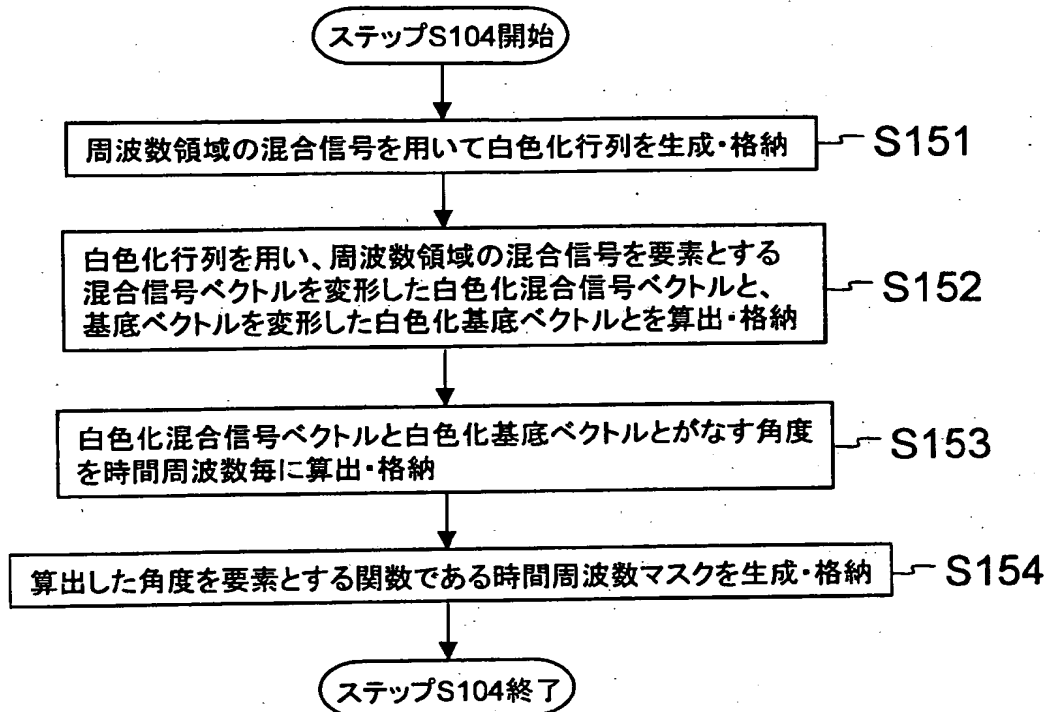


図23

[図24]

図24A

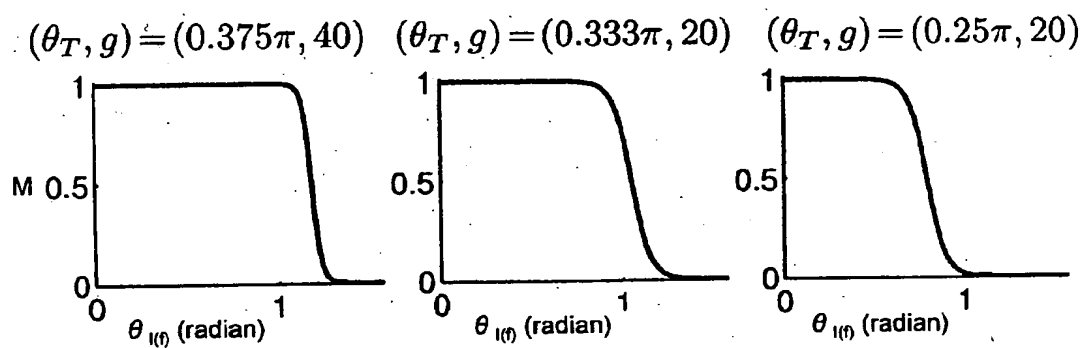
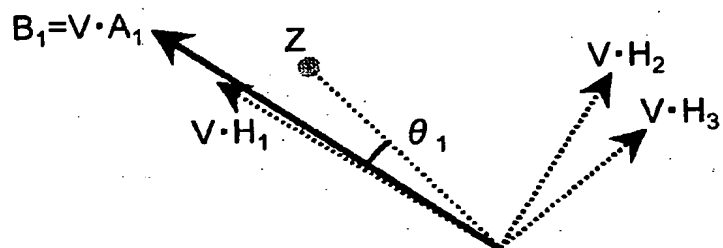
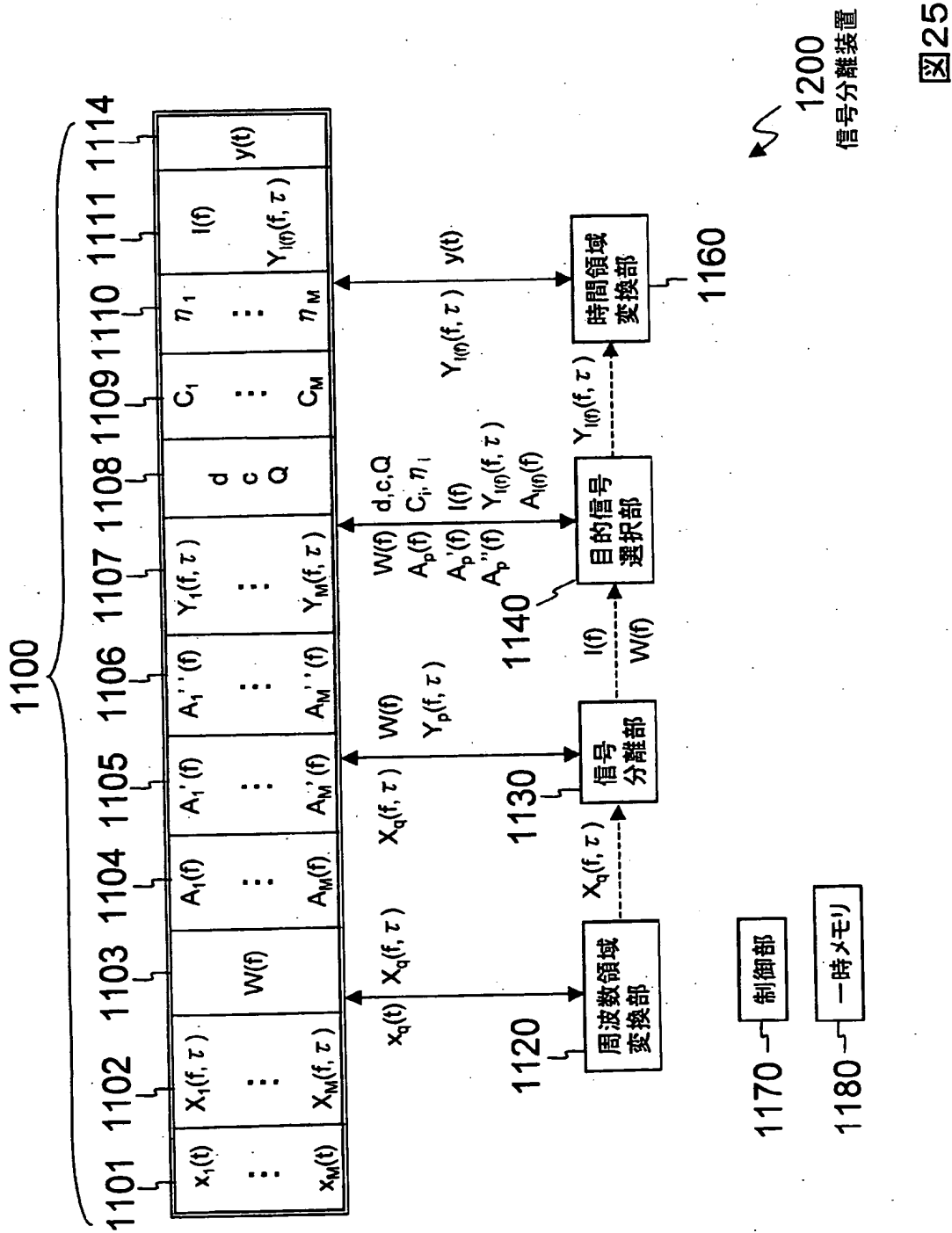


図24B



[図25]



[図26]

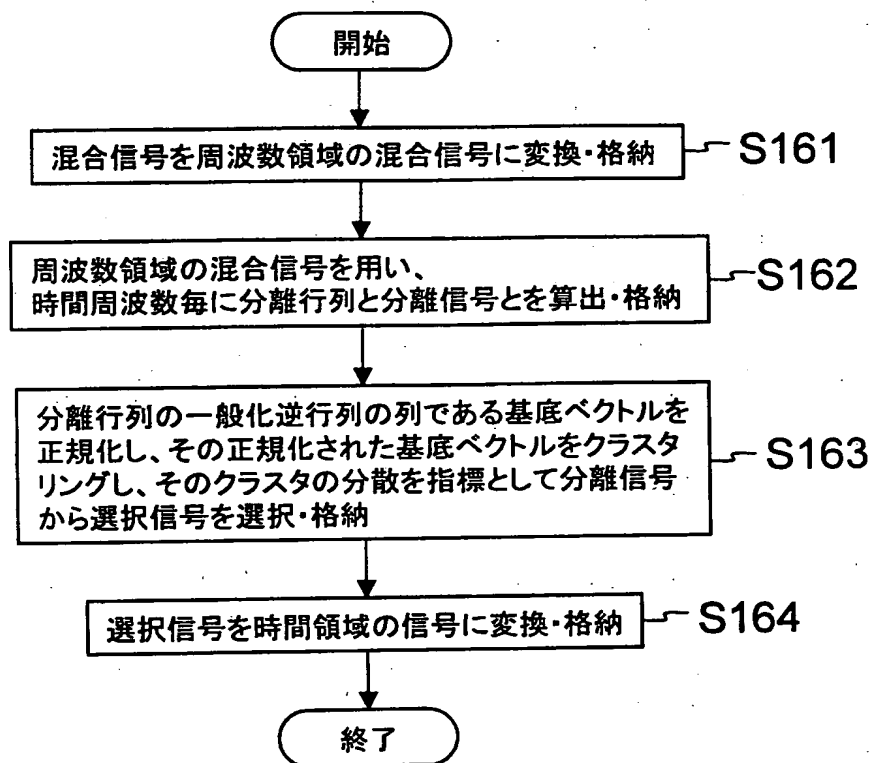
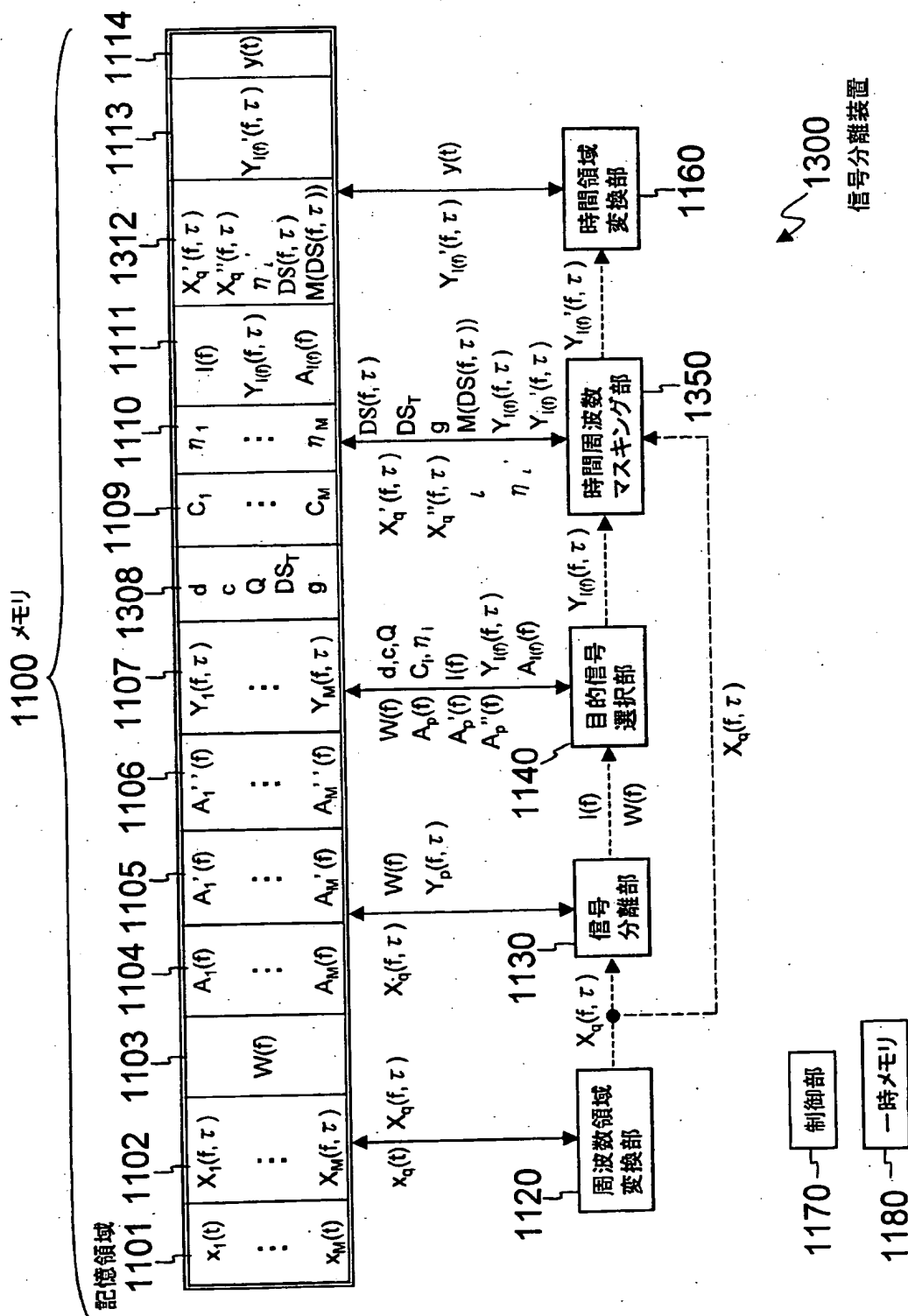


図26

[図27]



[図28]

図28A

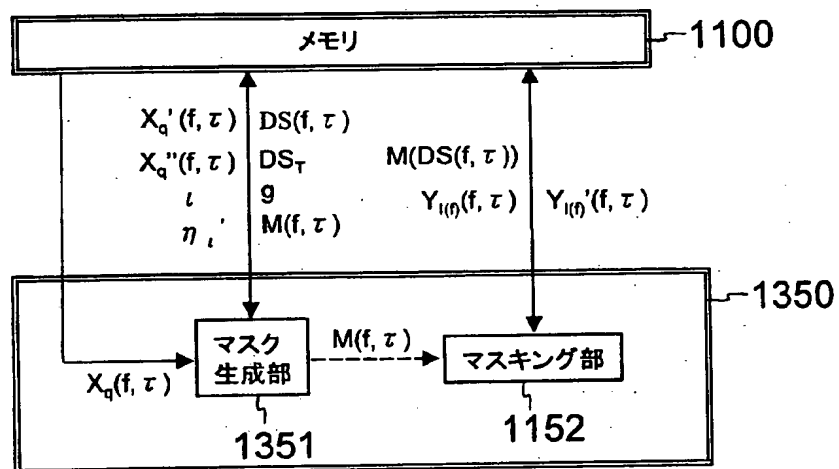
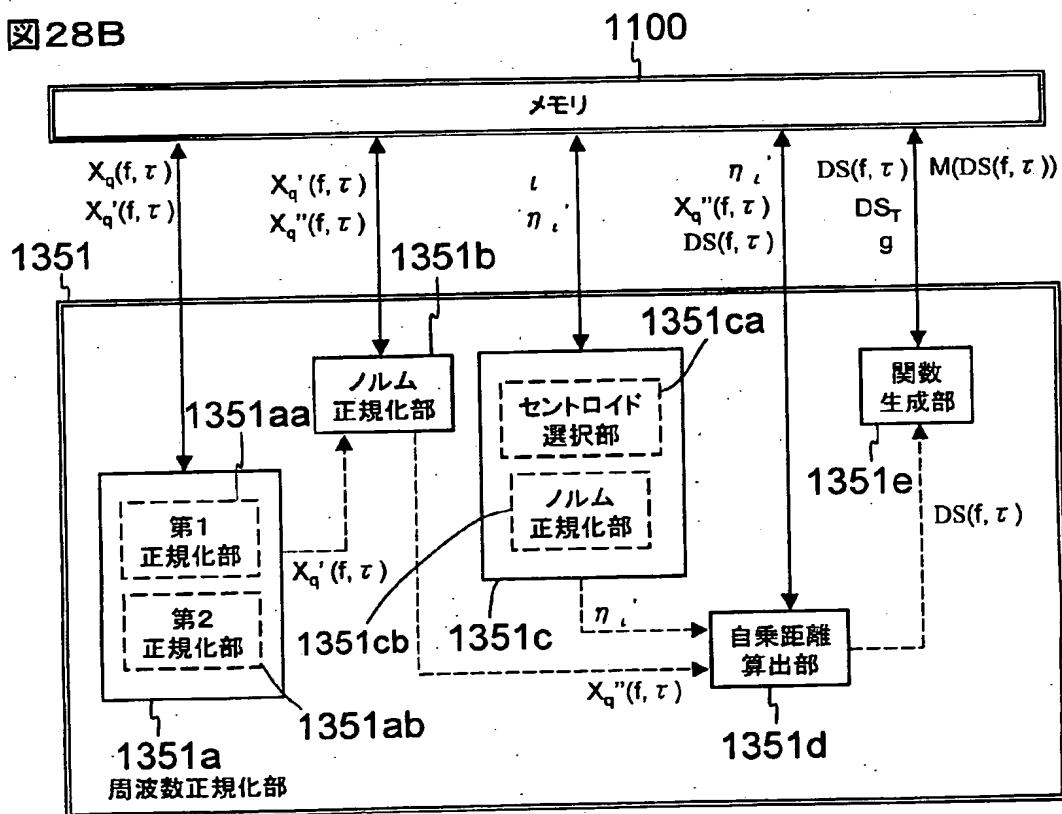


図28B



[図29]

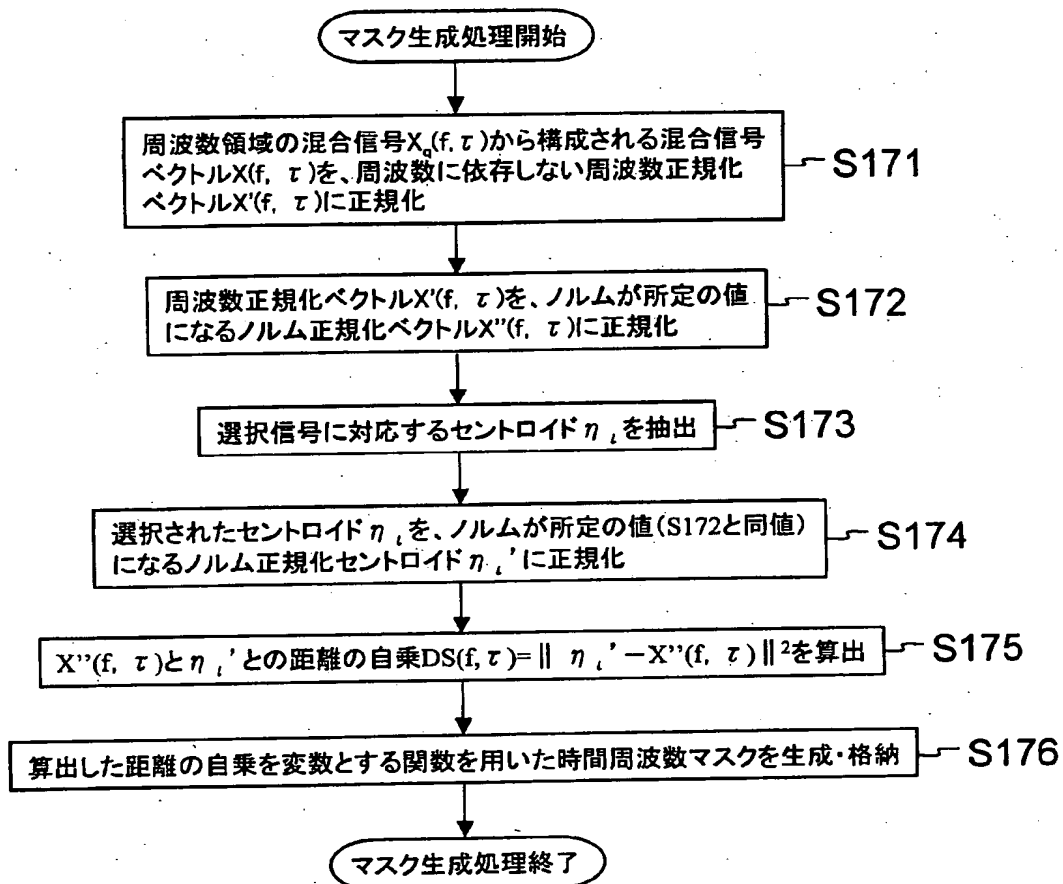


図29

[図30]

図30A

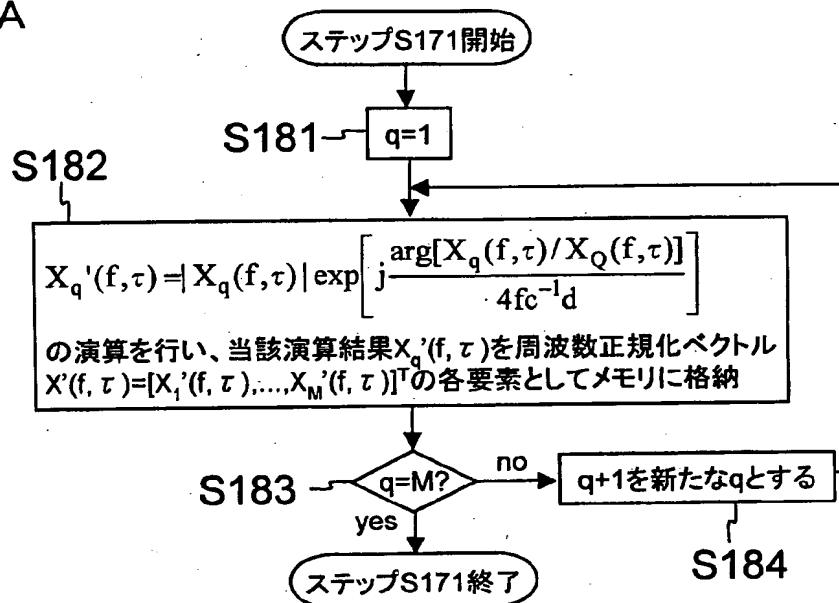
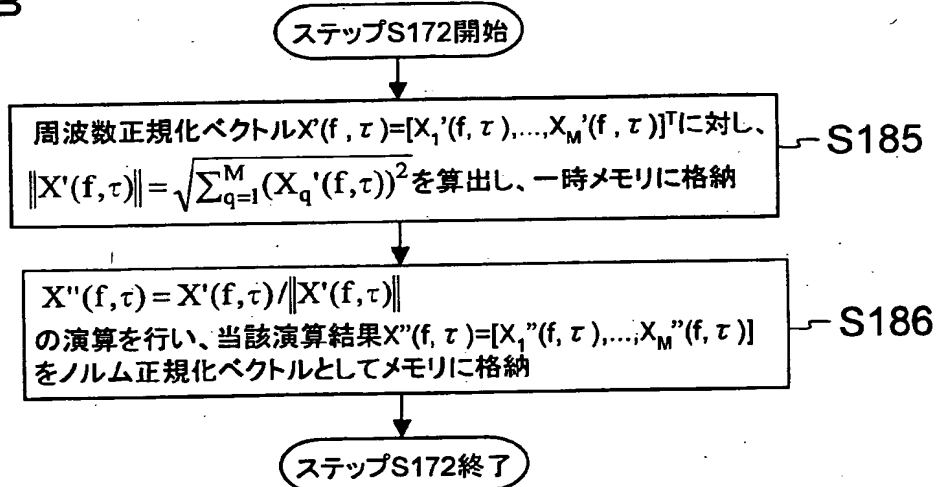


図30B



[図31]

図31A

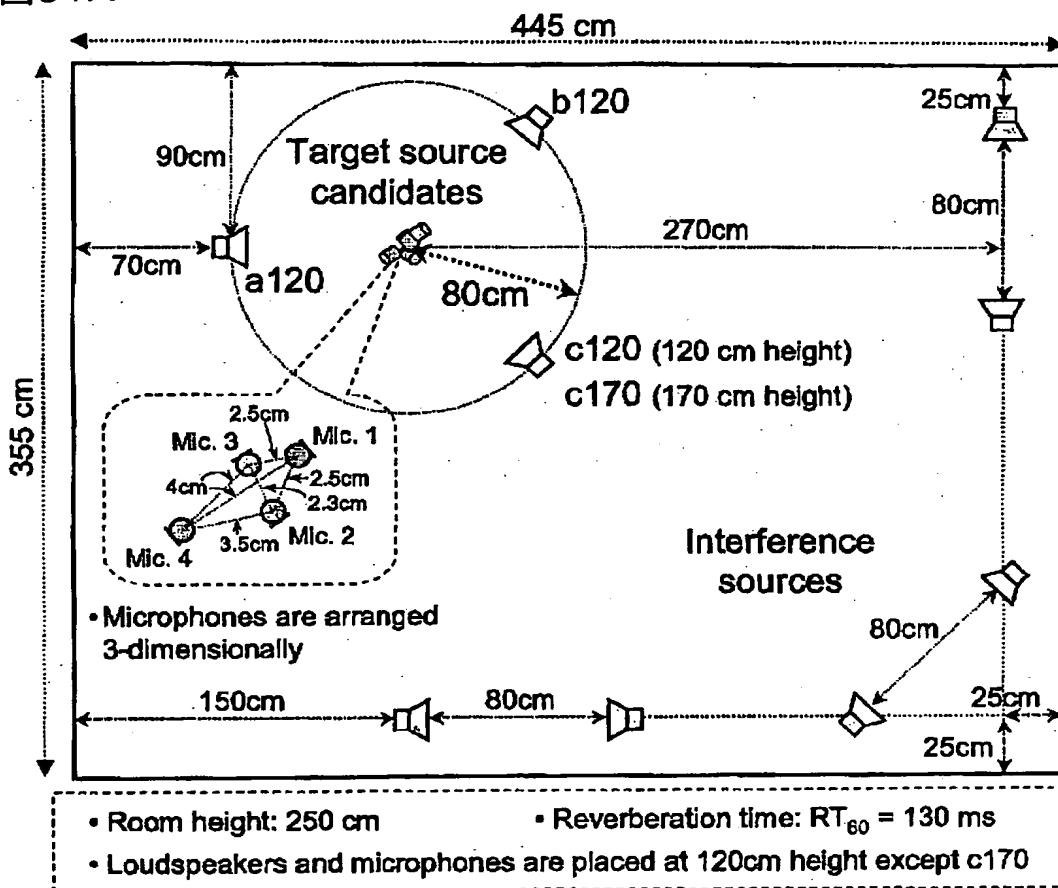


図31B

目的音声位置	a120	b120	c120	c170
ICA のみ	11.7 dB	11.7 dB	9.0 dB	12.9 dB
マスキングを併用 ($0.375\pi, 40$)	15.5 dB	14.3 dB	12.4 dB	16.7 dB
マスキングを併用 ($0.333\pi, 20$)	16.9 dB	15.4 dB	14.0 dB	18.1 dB
マスキングを併用 ($0.25\pi, 20$)	19.7 dB	17.9 dB	16.9 dB	20.8 dB

[図32]

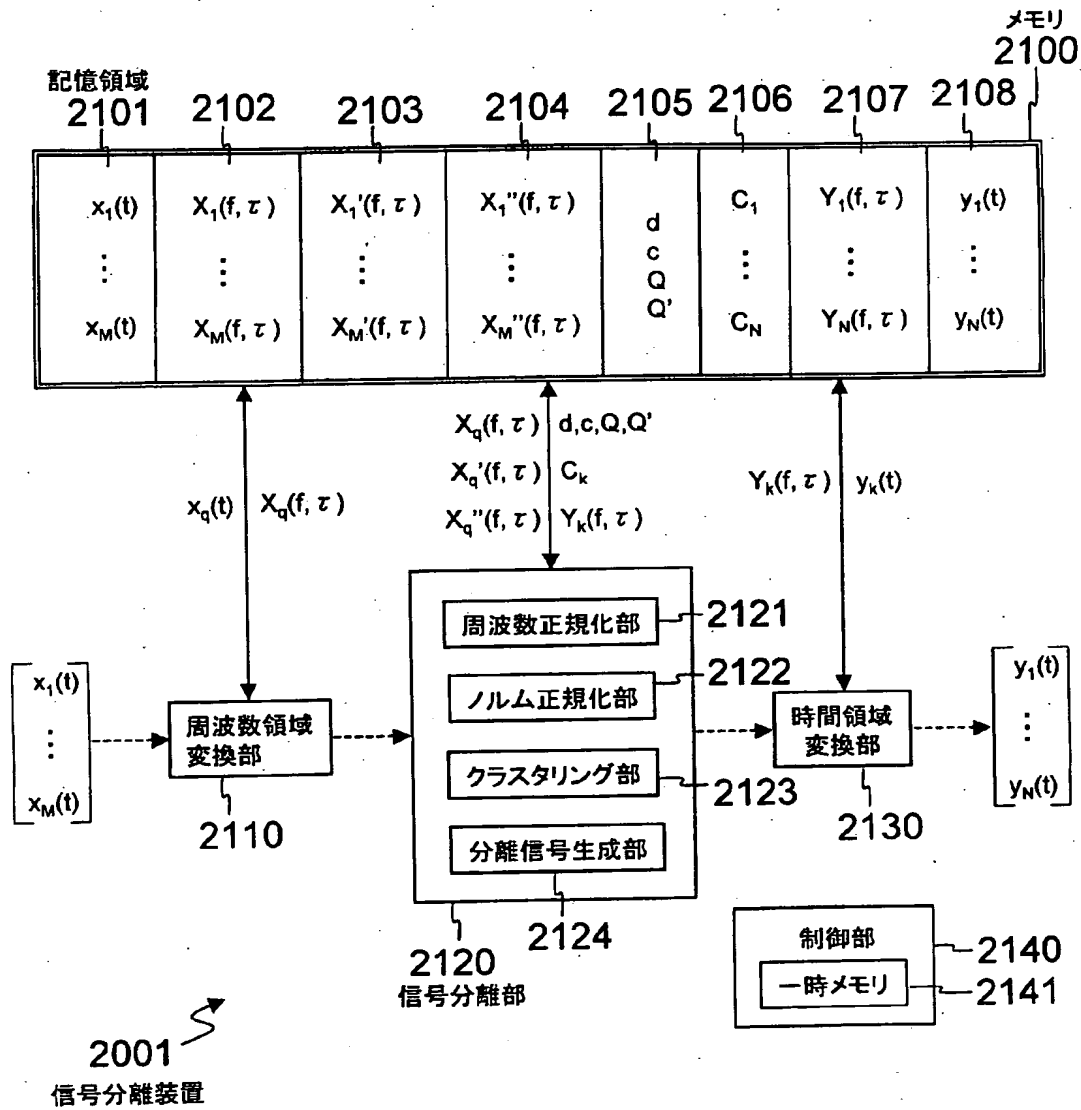


図32

[図33]

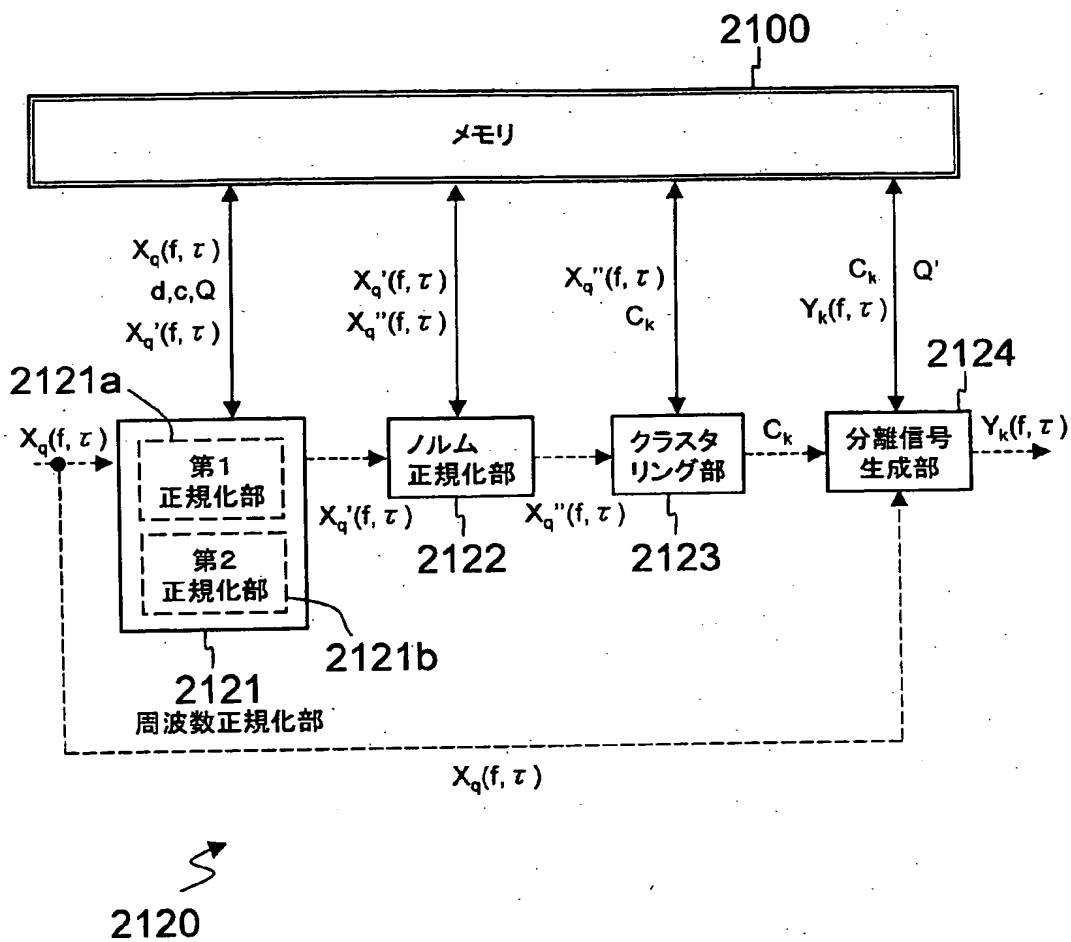


図33

[図34]

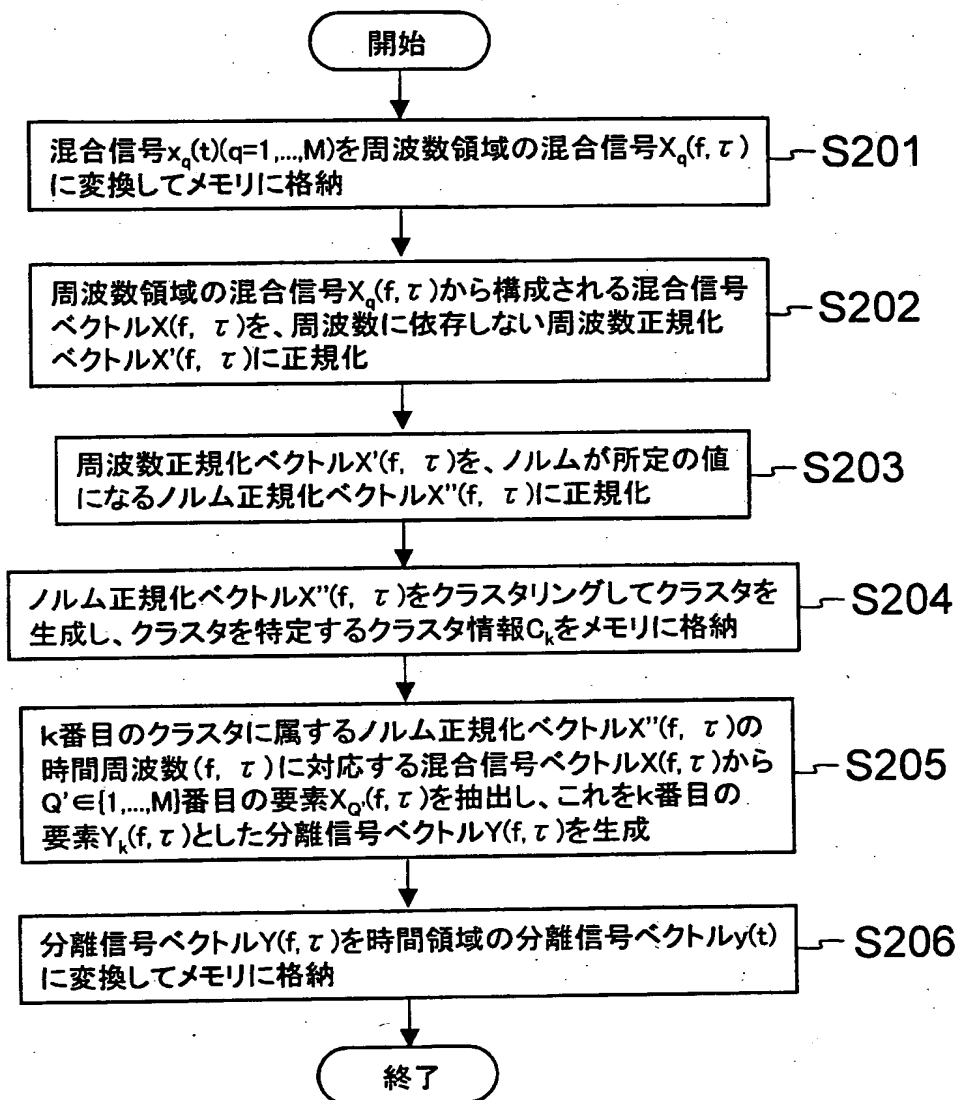


図34

[図35]

図35A

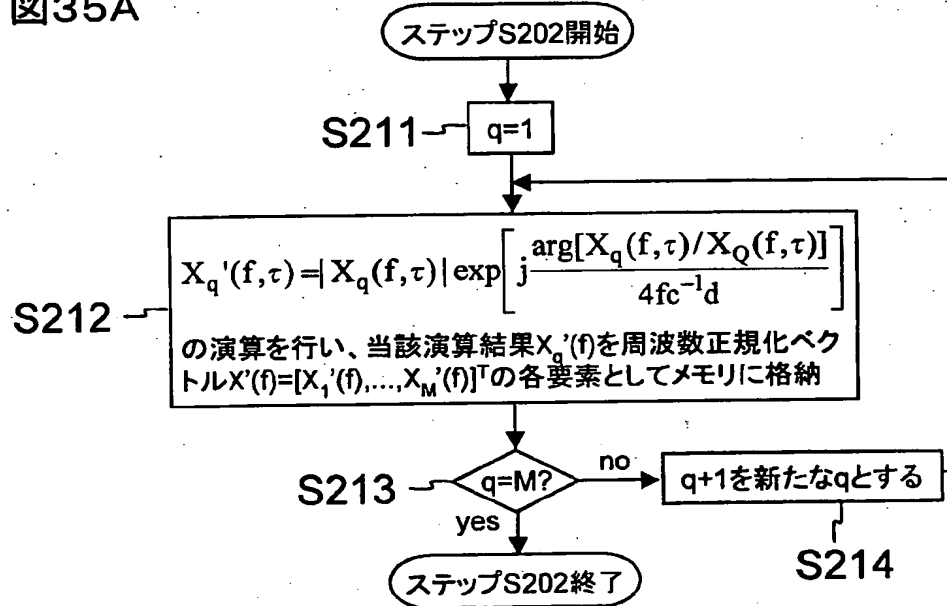
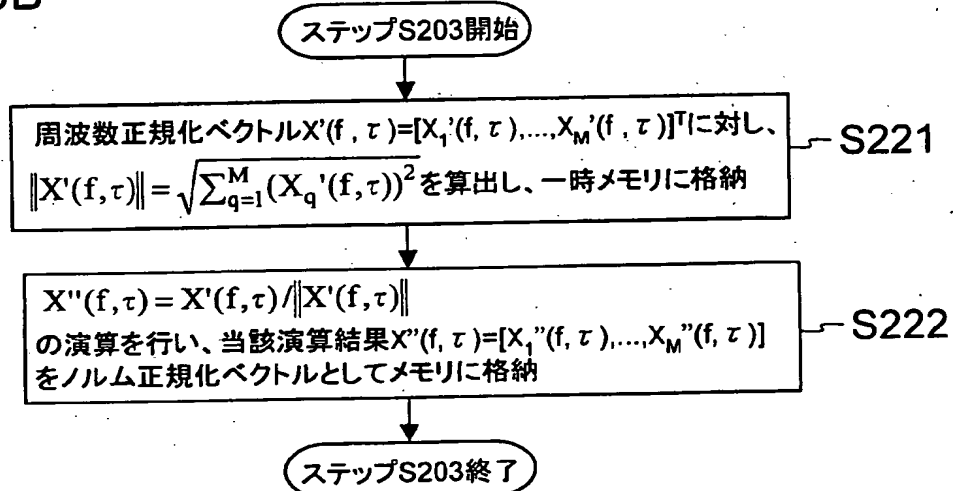


図35B



[図36]

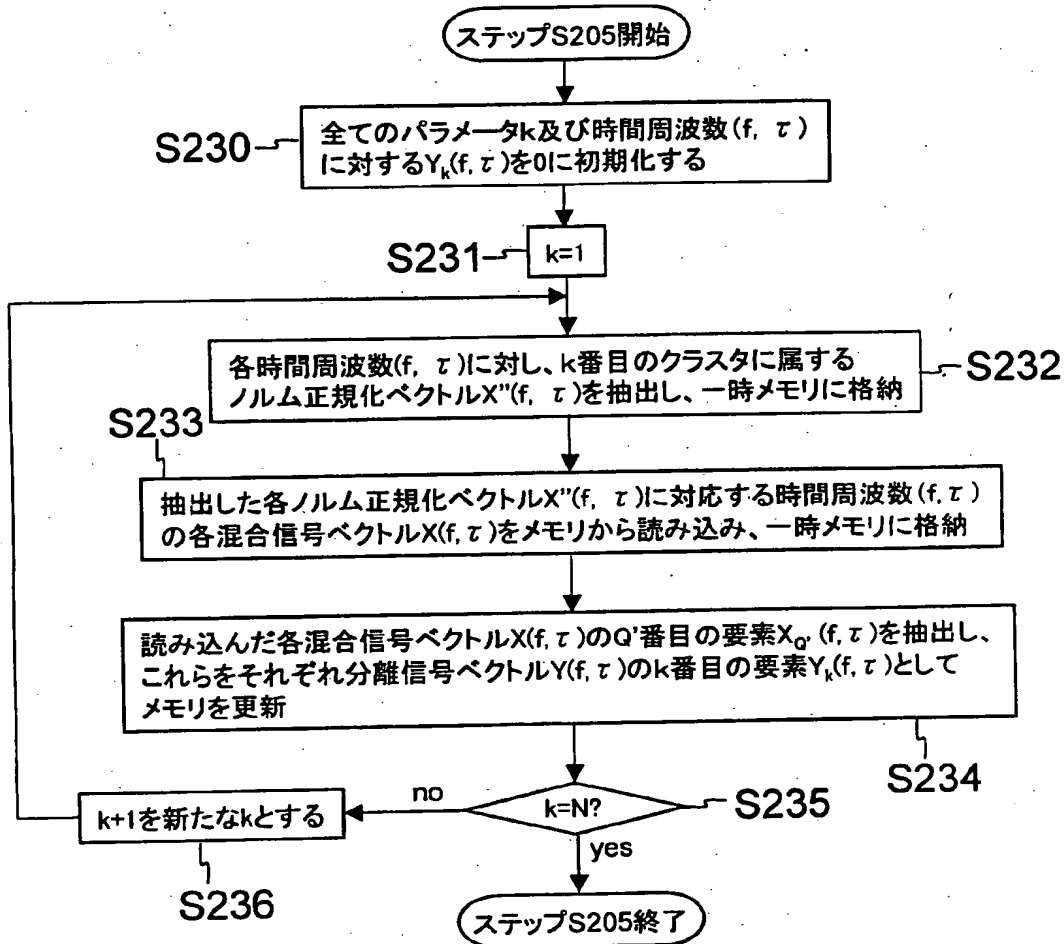


図36

[図37]

図37A

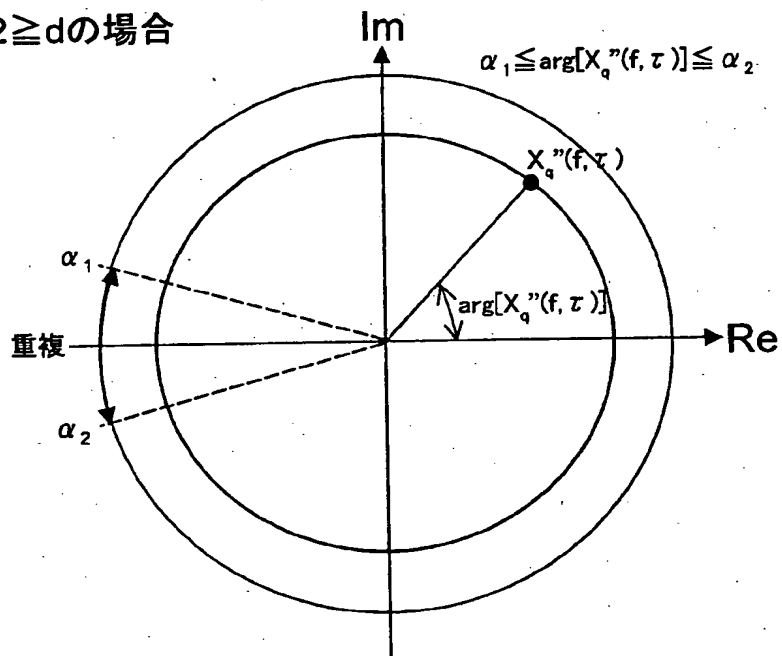
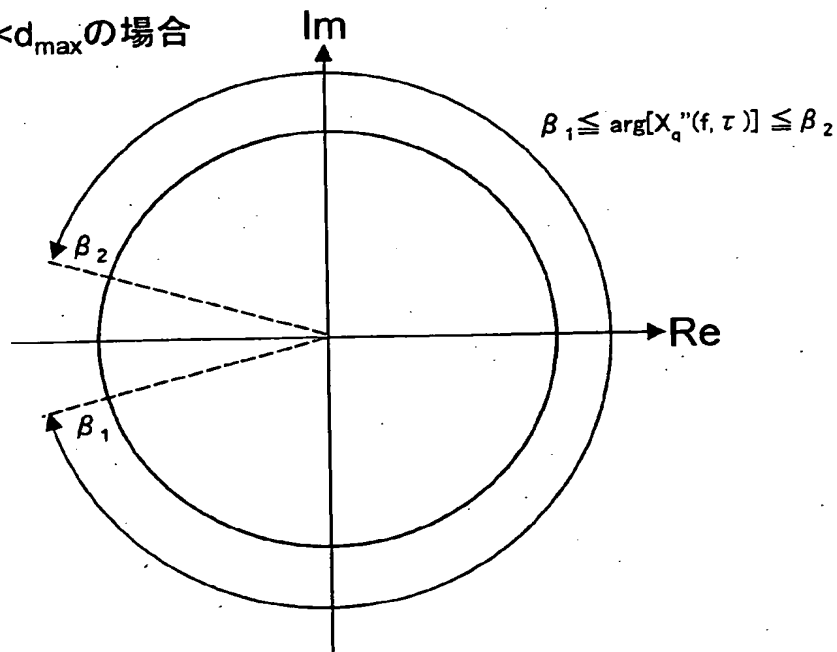
 $d_{\max}/2 \geq d$ の場合

図37B

 $d_{\max}/2 < d < d_{\max}$ の場合

[図38]

図38A

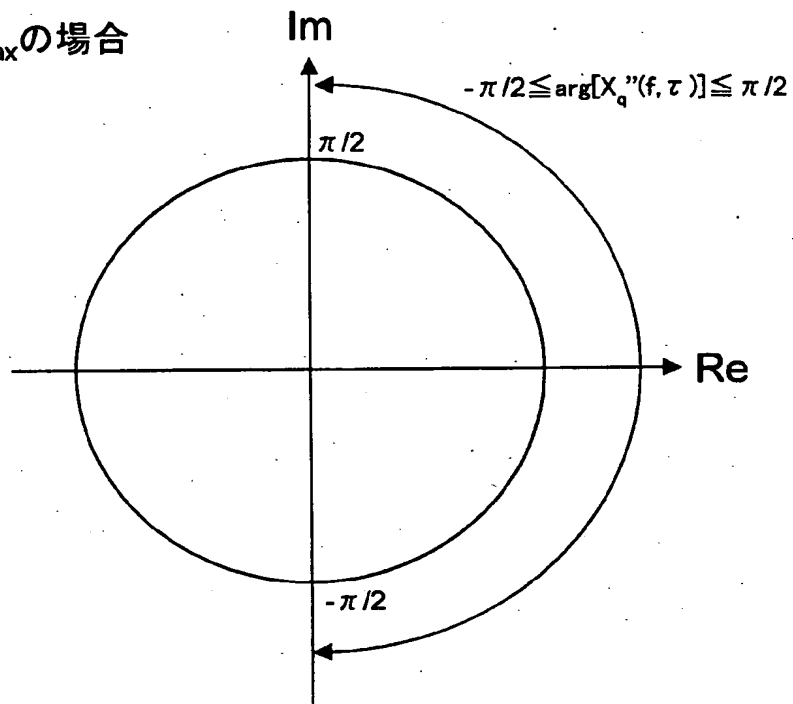
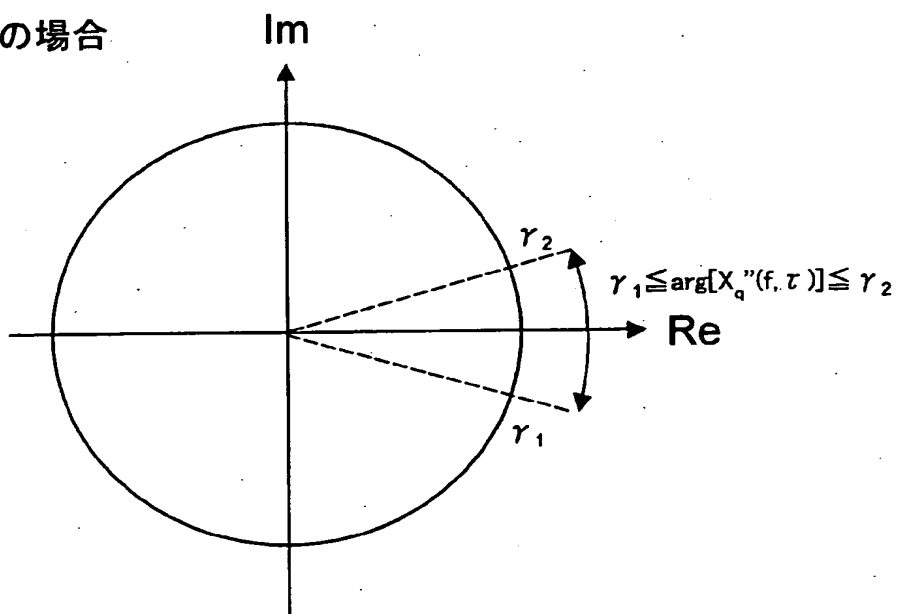
 $d=d_{\max}$ の場合

図38B

 $d > d_{\max}$ の場合

[図39]

図39A

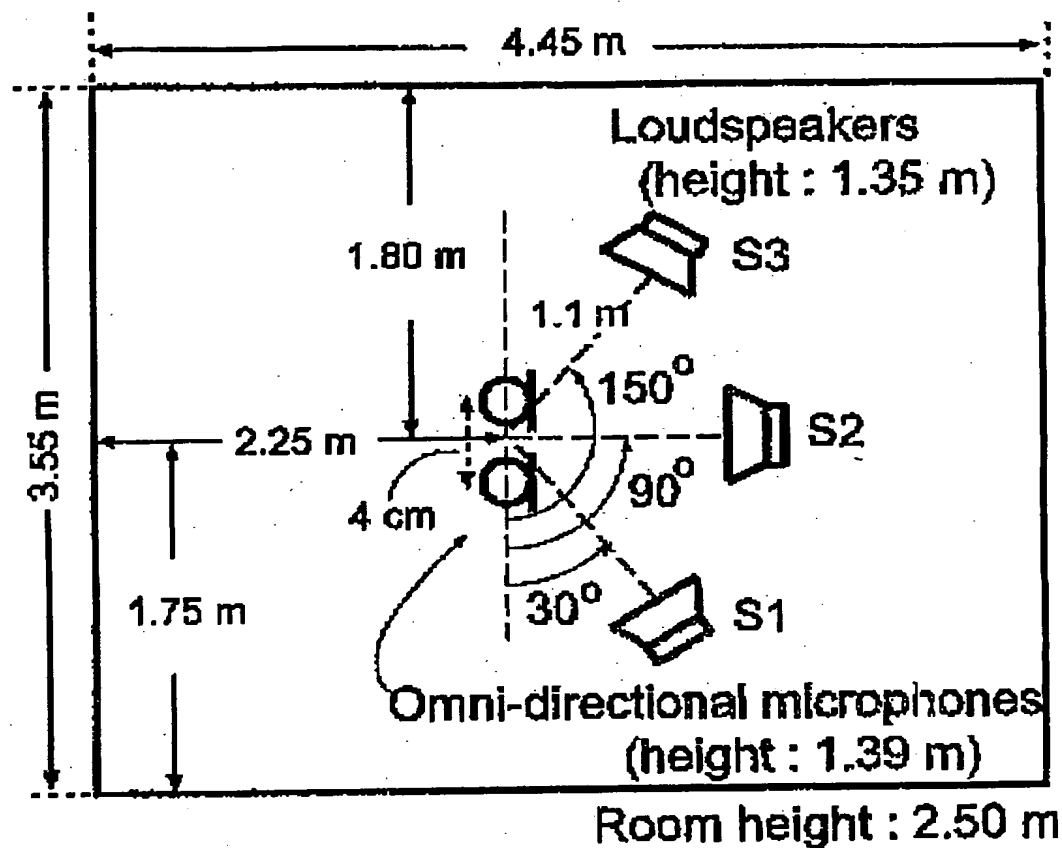


図39B

実験結果 (in[dB])

	SIR_1	SIR_2	SIR_3
InputSIR	-5.2	-1.5	-2.9
DOA (Previous)	17.7	6.3	12.7
Normalized obser. vector (Proposed)	16.4	6.0	13.0

[図40]

図40A

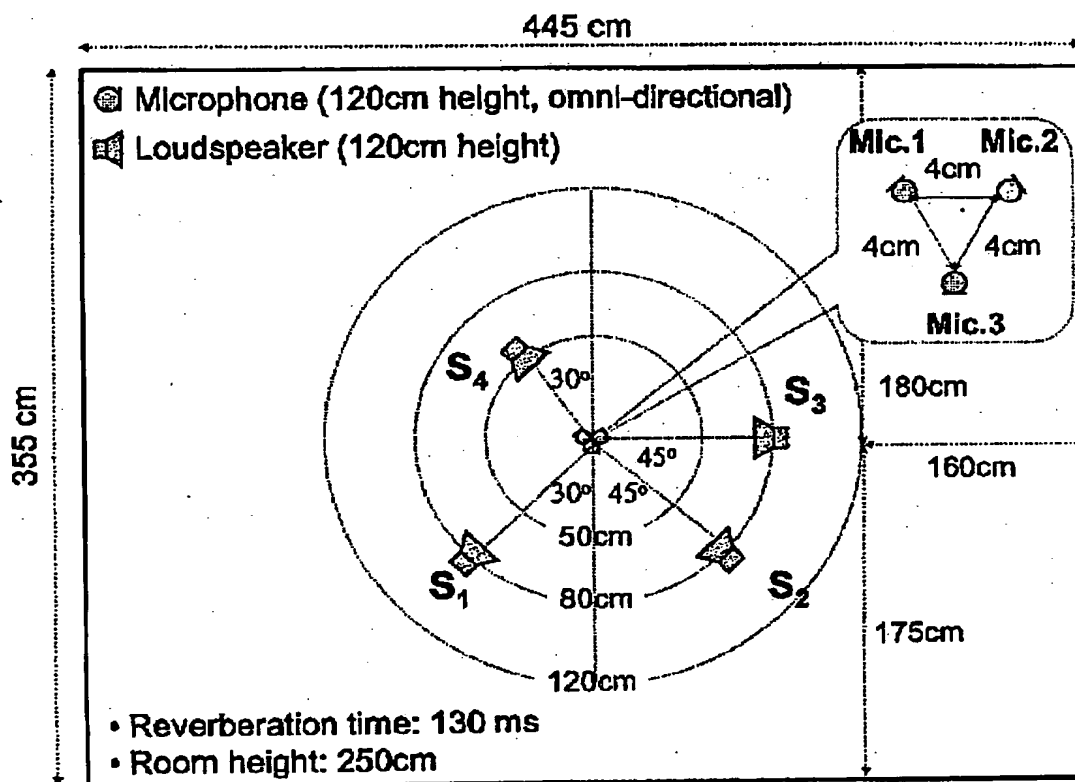


図40B

実験結果 (in[dB])

	SIR ₁	SIR ₂	SIR ₃	SIR ₄
InputSIR	-8.1	-6.0	-5.9	-0.8
Proposed	16.8	10.6	14.5	10.8

[図41]

図41A

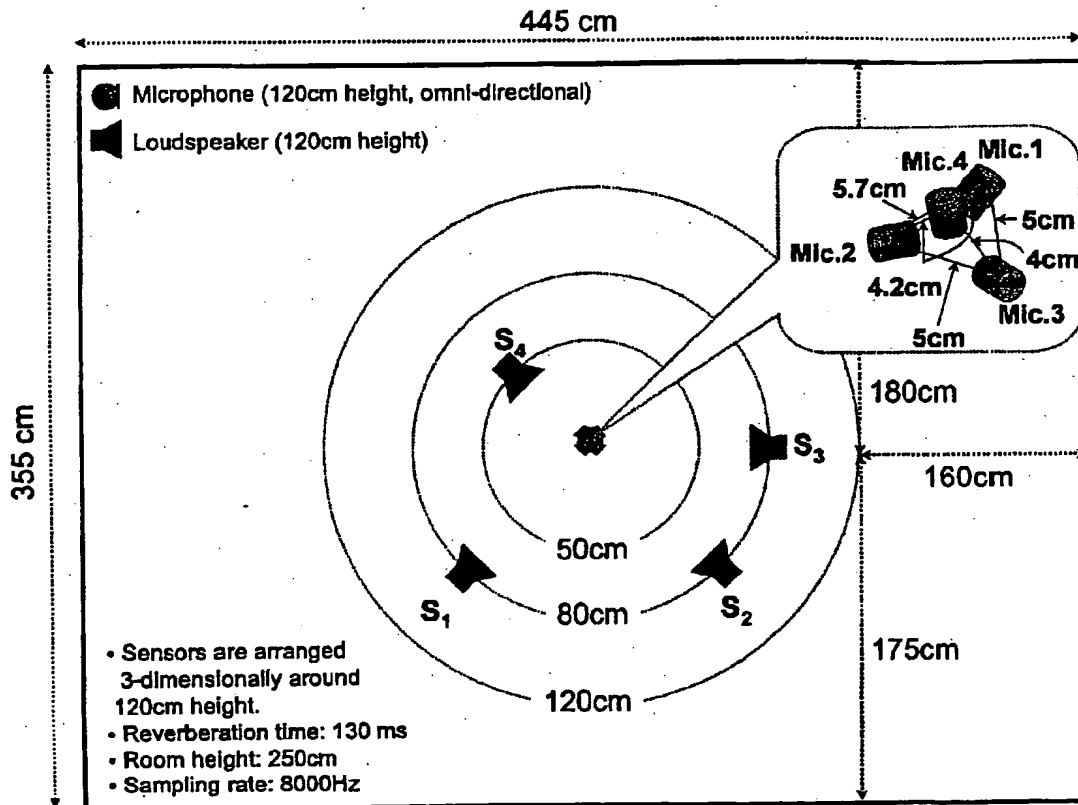


図41B

• 実験結果 (in [dB])

	SIR ₁	SIR ₂	SIR ₃	SIR ₄
InputSIR	-8.1	-5.3	-6.6	-0.6
Proposed	17.8	15.6	9.6	15.7

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2006/302092

A. CLASSIFICATION OF SUBJECT MATTER
G10L21/02 (2006.01)

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G10L21/02 (2006.01)

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
Jitsuyo Shinan Koho 1922-1996 Jitsuyo Shinan Toroku Koho 1996-2006
Kokai Jitsuyo Shinan Koho 1971-2006 Toroku Jitsuyo Shinan Koho 1994-2006

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
JSTPlus (JDream2), IEEE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2004-170552 A (Fujitsu Ltd.), 17 June, 2004 (17.06.04), Par. Nos. [0045] to [0054]; Fig. 6 (Family: none)	1-32
A	Sawada et al., "A robust and precise method for solving the permutation problem of frequency- domain blind source separation", IEEE Transactions on Speech and Audio Processing, 2004.09, Vol.12, No.5, pages 530 to 538	1-32
A	Sawada et al., "Polar coordinate based nonlinear function for frequency-domain source separation", IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, 01 March, 2003 (01.03.03), Vol.E86-A, Vol.3, pages 590 to 596	1-32

☒ Further documents are listed in the continuation of Box C.

☐ See patent family annex.

* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier application or patent but published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&" document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search
11 April, 2006 (11.04.06)

Date of mailing of the international search report
18 April, 2006 (18.04.06)

Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2006/302092

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
T	SAWADA et al., 'Taongen ni Taisuru Shuhasu Ryoiki Blind Ongen Bunri', The Japanese Society for Artificial Intelligence AI Challenge Symposium (Dai 22 Kai), 14 October, 2005 (14.10.05), pages 17 to 22	1-32
T	Sawada et al., "Blind extraction of a dominant source signal from mixtures of many sources", Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005 (ICASSP '05), 18 March, 2003 (18.03.03), Vol.3, pages 61 to 64	1-32